(19)日本国特許庁(JP)

# 再公表特許(A1)

(11)国際公開番号
WO2003/100422

発行日　平成17年9月29日(2005.9.29)

(43)国際公開日　平成15年12月4日(2003.12.4)

| (51)Int.Cl.[7] | F I | | |
|---|---|---|---|
| G01N 33/53 | G01N | 33/53 | M |
| C12M 1/00 | C12M | 1/00 | A |
| C12N 15/09 | C12Q | 1/68 | A |
| C12Q 1/68 | G01N | 33/566 | |
| G01N 33/566 | G01N | 37/00 | 102 |

審査請求 有　予備審査請求 未請求　(全 83 頁)　最終頁に続く

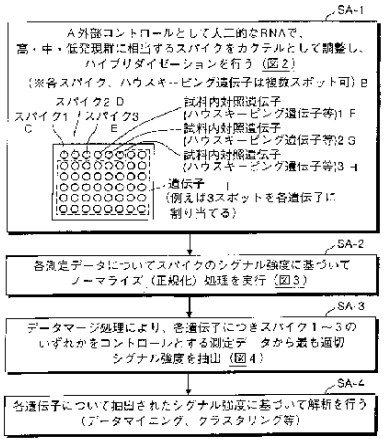
| | |
|---|---|
| 出願番号 | 特願2004-507828 (P2004-507828) |
| (21)国際出願番号 | PCT/JP2003/006677 |
| (22)国際出願日 | 平成15年5月28日(2003.5.28) |
| (31)優先権主張番号 | 特願2002-154773 (P2002-154773) |
| (32)優先日 | 平成14年5月28日(2002.5.28) |
| (33)優先権主張国 | 日本国(JP) |
| (81)指定国 | AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HU,IE,IT,LU,MC,NL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NI,NO,NZ,OM,PH,PL,PT,RO,RU,SC,SD,SE,SG,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,YU,ZA,ZM,ZW |

(71)出願人　501029744
株式会社ジェー・ジー・エス
東京都八王子市小宮町５１番地
(74)代理人　100089118
弁理士　酒井　宏明
(74)代理人　100113103
弁理士　香島　拓也
(72)発明者　安藤　賢
東京都八王子市小宮町５１　株式会社ジェー・ジー・エス内
(72)発明者　稲葉　二朗
東京都八王子市小宮町５１　株式会社ジェー・ジー・エス内

最終頁に続く

(54)【発明の名称】ＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体

(57)【要約】
本発明にかかるシステムは、複数のスパイク用遺伝子等から調整したＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のスパイクＲＮＡ等のシグナル強度に基づいてスキャニング画像の上記測定対象のＲＮＡの各スポットに対してシグナル強度のノーマライゼーションを行う。




A...PREPARING COCKTAIL OF SPIKES CORRESPONDING TO HIGH, MEDIUM- AND LOW EXPRESSION GROUPS WITH ARTIFICIAL RNAS AS EXTERNAL CONTROLS AND EFFECTING HYBRIDIZATION (FIG. 2)
B...(*:PLURAL SPOTS MAY BE FORMED FOR EACH SPIKE AND HOUSEKEEPING GENE)
C...SPIKE 1
D...SPIKE 2
E...SPIKE 3
F...CONTROL GENE IN SAMPLE (HOUSEKEEPING GENE, ETC.) 1
G...CONTROL GENE IN SAMPLE (HOUSEKEEPING GENE, ETC.) 2
H...CONTROL GENE IN SAMPLE (HOUSEKEEPING GENE, ETC.) 3
I...GENES (FOR EXAMPLE, ASSIGNING 3 SPOTS FOR EACH GENE)
J...EFFECTING NORMALIZATION BASED ON SIGNAL INTENSITIES OF SPIKES OF EACH MEASUREMENT DATA (FIG. 3)
K...EXTRACTING THE OPTIMUM SIGNAL INTENSITY FROM MEASUREMENT DATA USING ONE OF THE SPOTS 1 TO 3 OF EACH GENE VIA DATA MERGE PROCESSING (FIG. 4)
L...ANALYZING BASED ON SIGNAL INTENSITY EXTRACTED FOR EACH GENE (DATA MINING, CLUSTERING, ETC.)

【特許請求の範囲】
【請求項１】
複数の遺伝子から調整したコントロール用ＲＮＡおよび／または人工的に合成した複数のコントロール用ＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度に基づいてスキャニング画像の上記測定対象のＲＮＡの各スポットに対してシグナル強度のノーマライゼーションを行うこと、
を特徴とするＤＮＡマイクロアレイデータ処理方法。
【請求項２】
測定対象のＲＮＡに複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロール用ＲＮＡを混入して調整した試料に対してＤＮＡマイクロアレイを用いてハイブリダイゼーションを行うハイブリダイゼーションステップと、
上記ハイブリダイゼージョンステップにてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングするスキャニングステップと、
上記スキャニングステップにてスキャニングされたスキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズステップと、
上記ノーマライズステップにて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップと、
を含むことを特徴とするＤＮＡマイクロアレイデータ処理方法。
【請求項３】
上記スキャニングステップは、上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用ＲＮＡ毎に上記発光支持体をスキャニングすること、
を特徴とする請求の範囲第２項に記載のＤＮＡマイクロアレイデータ処理方法。
【請求項４】
上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない、複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第１項から第３項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項５】
上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または、非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合で、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない、複数の遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第１項から第３項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項６】
上記複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いること、
を特徴とする請求の範囲第１項から第５項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項７】
上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、

バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整すること、
を特徴とする請求の範囲第１項から第６項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項８】
測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のコントロール用スポットに結合したＲＮＡのシグナル強度に基づいてスキャニング画像の測定対象のＲＮＡが結合した各スポットに対してシグナル強度のノーマライゼーションを行うこと、
を特徴とするＤＮＡマイクロアレイデータ処理方法。
【請求項９】
測定対象遺伝子のＲＮＡに対して、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行うハイブリダイゼーションステップと、
上記ハイブリダイゼージョンステップにてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングするスキャニングステップと、
上記スキャニングステップにてスキャニングされたスキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズステップと、
上記ノーマライズステップにて各コントロール用スポットによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップと、
を含むことを特徴とするＤＮＡマイクロアレイデータ処理方法。
【請求項１０】
上記スキャニングステップは、上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎に上記発光支持体をスキャニングすること、
を特徴とする請求の範囲第９項に記載のＤＮＡマイクロアレイデータ処理方法。
【請求項１１】
上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いること、
を特徴とする請求の範囲第８項から第１０項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項１２】
上記ノーマライズステップは、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化ステップ、
をさらに含むことを特徴とする請求の範囲第２項から第７項、第９項から第１１項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項１３】
上記ノーマライズステップは、上記統合化ステップにて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポ

ット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定ステップと、
上記棄却値選定ステップにより選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定ステップと、
上記代表値決定ステップにて決定された上記代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定する回帰式補正ステップと、
をさらに含むことを特徴とする請求の範囲第１２項に記載のＤＮＡマイクロアレイデータ処理方法。
【請求項１４】
上記ノーマライズステップは、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正ステップ、
をさらに含むことを特徴とする請求の範囲第２項から第７項、第９項から第１３項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理方法。
【請求項１５】
ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置と、当該スキャナ装置によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置とを含んで構成されるＤＮＡマイクロアレイデータ処理システムであって、
上記スキャナ装置は、
複数の遺伝子から調整したコントロール用ＲＮＡおよび／または人工的に合成したコントロール用ＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャニング手段、
を備え、
上記ＤＮＡマイクロアレイデータ処理装置は、
上記スキャニング手段にてスキャニングされたスキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズ手段と、
上記ノーマライズ手段にて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段と、
を備えることを特徴とするＤＮＡマイクロアレイデータ処理システム。
【請求項１６】
上記スキャニング手段は、
上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように感度を調整して上記コントロール用ＲＮＡ毎に上記発光支持体をスキャニングするスキャニング感度調整手段、
をさらに備えることを特徴とする請求の範囲第１５項に記載のＤＮＡマイクロアレイデータ処理システム。
【請求項１７】
上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない、複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第１５項または第１６項に記載のＤＮＡマイクロアレイデータ処理システム。
【請求項１８】

上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第１５項または第１６項に記載のＤＮＡマイクロアレイデータ処理システム。

【請求項１９】
上記複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いること、
を特徴とする請求の範囲第１５項から第１８項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理システム。

【請求項２０】
上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整すること、
を特徴とする請求の範囲第１５項から第１９項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理システム。

【請求項２１】
ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置と、当該スキャナ装置によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置とを含んで構成されるＤＮＡマイクロアレイデータ処理システムであって、
上記スキャナ装置は、
測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャニング手段、
を備え、
上記ＤＮＡマイクロアレイデータ処理装置は、
上記スキャニング手段にてスキャニングされたスキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポット遺伝子毎に実行するノーマライズ手段と、
上記ノーマライズ手段にて各コントロール用スポットによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段と、
を備えることを特徴とするＤＮＡマイクロアレイデータ処理システム。

【請求項２２】
上記スキャニング手段は、
上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように感度を調整して上記コントロール用スポットの遺伝子毎に上記発光支持体をスキャニングするスキャニング感度調整手段、
をさらに備えることを特徴とする請求の範囲第２１項に記載のＤＮＡマイクロアレイデータ処理システム。

【請求項２３】
上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いること、
を特徴とする請求の範囲第２１項または第２２項に記載のＤＮＡマイクロアレイデータ処理システム。

【請求項２４】
上記ノーマライズ手段は、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化手段、
をさらに含むことを特徴とする請求の範囲第１５項から第２３項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理システム。
【請求項２５】
上記ノーマライズ手段は、上記統合化手段にて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定手段と、
上記棄却値選定手段により選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定手段と、
上記仮代表値決定手段にて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正する回帰式補正手段と、
をさらに含むことを特徴とする請求の範囲第２４項に記載のＤＮＡマイクロアレイデータ処理システム。
【請求項２６】
上記ノーマライズ手段は、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正手段、
をさらに含むことを特徴とする請求の範囲第１５項から第２５項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理システム。
【請求項２７】
複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロールＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得手段と、
上記スキャニング画像取得手段にて取得された上記スキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズ手段と、
上記ノーマライズ手段にて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段と、
を備えることを特徴とするＤＮＡマイクロアレイデータ処理装置。
【請求項２８】
上記スキャニング画像は、
上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用ＲＮＡ毎にスキャニングされた複数の画像であること、
を特徴とする請求の範囲第２７項に記載のＤＮＡマイクロアレイデータ処理装置。
【請求項２９】
上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第２７項または第２８項に記載のＤＮＡマイクロアレイデータ処

理装置。
【請求項３０】
上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第２７項または第２８項に記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３１】
上記複数のコントロール用ＲＮＡは、測定対象遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いること、
を特徴とする請求の範囲第２７項から第３０項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３２】
上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウイルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整すること、
を特徴とする請求の範囲第２７項から第３１項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３３】
測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得手段と、
上記スキャニング画像取得手段にて取得された上記スキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズ手段と、
上記ノーマライズ手段にて各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段と、
を備えることを特徴とするＤＮＡマイクロアレイデータ処理装置。
【請求項３４】
上記スキャニング画像は、
上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎にスキャニングされた複数の画像であること、
を特徴とする請求の範囲第３３項に記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３５】
上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いること、
を特徴とする請求の範囲第３３項または第３４項に記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３６】
上記ノーマライズ手段は、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化手段、
をさらに含むことを特徴とする請求の範囲第２７項から第３５項のいずれか一つに記載の

ＤＮＡマイクロアレイデータ処理装置。
【請求項３７】
上記ノーマライズ手段は、上記統合化手段にて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定手段と、上記棄却値選定手段により選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定手段と、
上記仮代表値決定手段にて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定する回帰式補正手段と、
をさらに含むことを特徴とする請求の範囲第３６項に記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３８】
上記ノーマライズ手段は、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正手段、
をさらに含むことを特徴とする請求の範囲第２７項から第３７項のいずれか一つに記載のＤＮＡマイクロアレイデータ処理装置。
【請求項３９】
複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロール用ＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得ステップと、
上記スキャニング画像取得ステップにて取得された上記スキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズステップと、
上記ノーマライズステップにて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップと、
を含むＤＮＡマイクロアレイデータ処理方法をコンピュータに実行させることを特徴とするプログラム。
【請求項４０】
上記スキャニング画像は、
上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用ＲＮＡ毎にスキャニングされた複数の画像であること、
を特徴とする請求の範囲第３９項に記載のプログラム。
【請求項４１】
上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、
を特徴とする請求の範囲第３９項または第４０項に記載のプログラム。
【請求項４２】
上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであること、

を特徴とする請求の範囲第３９項または第４０項に記載のプログラム。
【請求項４３】
上記複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いること、
を特徴とする請求の範囲第３９項から第４２項のいずれか一つに記載のプログラム。
【請求項４４】
上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整すること、
を特徴とする請求の範囲第３９項から第４３項のいずれか一つに記載のプログラム。
【請求項４５】
測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得ステップと、
上記スキャニング画像取得ステップにて取得された上記スキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズステップと、
上記ノーマライズステップにて各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップと、
を含むＤＮＡマイクロアレイデータ処理方法をコンピュータに実行させることを特徴とするプログラム。
【請求項４６】
上記スキャニング画像は、
上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎にスキャニングされた複数の画像であること、
を特徴とする請求の範囲第４５項に記載のプログラム。
【請求項４７】
上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いること、
を特徴とする請求の範囲第４５項または第４６項に記載のプログラム。
【請求項４８】
上記ノーマライズステップは、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化ステップ、
をさらに含むことを特徴とする請求の範囲第３９項から第４７項のいずれか一つに記載のプログラム。
【請求項４９】
上記ノーマライズステップは、上記統合化ステップにて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定ステップと、
上記棄却値選定ステップにより選定された上記棄却値を用いて棄却した後の残された同一

遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定ステップと、
上記仮代表値決定ステップにて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定する回帰式補正ステップと、
をさらに含むことを特徴とする請求の範囲第４８項に記載のプログラム。
【請求項５０】
上記ノーマライズステップは、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正ステップ、
をさらに含むことを特徴とする請求の範囲第３９項から第４９項のいずれか一つに記載のプログラム。
【請求項５１】
上記請求の範囲第３９項から第５０項のいずれか一つに記載されたプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。
【発明の詳細な説明】
技術分野
本発明は、ＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体に関し、特に、例えば二蛍光標識法によるスパイク実験において複数のスパイク等を用いて効率的に発現画像データの解析を行うＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体に関する。
背景技術
ＤＮＡマイクロアレイ（本明細書においてＤＮＡチップなどを含んだ概念で使用する。）は、ガラスなどの支持体上に遺伝子ＤＮＡ及びオリゴヌクレオチドなどを高密度で配列したデバイスであり、ｃＤＮＡ、ｃＲＮＡあるいはゲノムＤＮＡなどとハイブリダイゼーションを行うことにより、遺伝子の網羅的な解析をすることができる。すなわちＤＮＡマイクロアレイにより、発生・分化における遺伝子発現プロファイリング、病態変動遺伝子群の解析及び遺伝子多型解析などが可能であることが知られている（Ｓｃｈｅｎａ，Ｍｅｔ ａｌ．：Ｑｕａｎｔｉｔａｔｉｖｅ ｍｏｎｉｔｏｒｉｎｇ ｏｆ ｇｅｎｅ ｅｘｐｒｅｓｓｉｏｎ ｐａｔｔｅｒｎｓ ｗｉｔｈ ａ ｃｏｍｐｌｅｍｅｎｔａｒｙ ＤＮＡ ｍｉｃｒｏａｒｒａｙ，Ｓｃｉｅｎｃｅ ２７０，４６７－４７０，１９９５）。病態変動遺伝子群の解析として例えば、白血病の詳細分類を行う事で、治療法を変更すべき群を識別・鑑別しようとする応用例をあげることができる（Ｋｈａｎ，Ｊ ｅｔ ａｌ．：Ｃｌａｓｓｉｆｉｃａｔｉｏｎ ａｎｄ ｄｉａｇｎｏｓｔｉｃ ｐｒｅｄｉｃｔｉｏｎ ｏｆ ｃａｎｃｅｒｓ ｕｓｉｎｇ ｇｅｎｅ ｅｘｐｒｅｓｓｉｏｎ ｐｒｏｆｉｌｉｎｇ ａｎｄ ａｒｔｉｆｉｃｉａｌ ｎｅｕｒａｌ ｎｅｔｗｏｒｋｓ、Ｎａｔｕｒｅ Ｍｅｄｉｃｉｎｅ ７，６７３－６７９，（２００１））。これに基づき、数百の関連遺伝子の発現をＤＮＡマイクロアレイを用いて調べ、発現パターンをクラスター解析してより精度の高い診断を試みようとしている。
また、肝臓癌においてｌｕｃｉｆｅｒａｓｅ遺伝子をスパイク（コントロール）として用い、９３０種の特異的な遺伝子を搭載したＤＮＡマイクロアレイ解析により、遺伝子異常が解析された（Ｓｈｉｒｏｔａ，Ｙ ｅｔ ａｌ．：Ｉｄｅｎｔｉｆｉｃａｔｉｏｎ ｏｆ Ｄｉｆｆｅｒｅｎｔｉａｌｌｙ Ｅｘｐｒｅｓｓｅｄ Ｇｅｎｅｓ ｉｎ Ｈｅｐａｔｏｃｅｌｌｕｌａｒ Ｃａｒｃｉｎｏｍａ ｗｉｔｈ ｃＤＮＡ Ｍｉｃｒｏａｒｒａｙｓ，Ｈｅｐａｔｏｌｏｇｙ ３３，８３２－８４０，２００１）。
ＤＮＡマイクロアレイは上記のように、発現解析、病態解析など広範な分野で応用が期待される手法であるが、ごく最近の技術であるため、研究者によるそれらの検討はまだ途中

段階であり、確立した方法論はいまだ現れていない。一方、ＤＮＡマイクロアレイによる解析手法は、スライドガラスなどの支持体上に高密度に固定されたＤＮＡプローブなどと標識した試料核酸をハイブリダイスする過程、標識を検出する過程および検出データを解析する過程で構成されており、その各過程において、これまでに様々な問題が表出しており、その解決が検討されている。

それらの問題点とは、主に感度、ばらつき（同一マイクロアレイ内の同一遺伝子間におきるもの、またはマイクロアレイ間の誤差）であるが、それらに加え反応工程における誤差の積み重ね、測定機器や支持体等の特性によるバイアス、なども含まれる。

一方、ＤＮＡマイクロアレイによる解析方法が普及するにつれ、診断での応用が期待されている。そのためには（１）患者サンプル同士が比較可能であること、（２）疾患特有の遺伝子発現の変化を感知できること（＝低発現の遺伝子であってもその発現変化を感知できること）等の工夫が必要となる。

これらの問題の解決策（特に上記（１）の解決策）として、例えば同一マイクロアレイ内の同一遺伝子のばらつきやマイクロアレイ間の誤差を少なくするため、各種コントロールＲＮＡなどの内部標準物質を用いた二蛍光標識法を用いた競合的ハイブリダイゼーションを行い、更に検出データをコンピュータを用いて大量解析する手法により、測定の精度や再現性の向上が図られている。ここで、二蛍光標識法とは、二つのｍＲＮＡサンプルをそれぞれ異なる蛍光で標識し、同一マイクロアレイ上で競合的ハイブリダイゼーションを行って、両方の蛍光を測定・比較することで遺伝子発現の差を検出する手法である。そこでこの方法で一方を測定対象（検体ＲＮＡ）とし、一方を対照ＲＮＡとして競合させ、内部標準でノーマライズして比を取れば、対照ＲＮＡが一定でさえあれば、理論上は各遺伝子におけるそれぞれの比を、測定対象同士（＝異なるマイクロアレイ同士）で比較することができることになる。

現在最も新しく、かつ一般的教科書として利用されているＳｔｅｅｎ　Ｋｎｕｄｓｅｎ著「ＤＮＡマイクロアレイデータ解析入門」（羊土社、２００２年）の「第３章データ解析の基本」では、上述の内部標準によるノーマライズを「スケール化」と称して、「いわゆるメンテナンス遺伝子（ハウスキーピング遺伝子）を含めることで可能となる。」と述べている。

しかし、問題はそれほど単純ではない。二蛍光標識法において、一方を検体ＲＮＡ、一方を対照ＲＮＡとした場合に、何を対照ＲＮＡとするかの選択自体がまず問題となってくる。例えば、臨床サンプルの個々の検証では検体を癌組織、対照を（癌組織提供者の）正常組織、というように設定することが可能であるが、実際の臨床で個々人に最適な対照群を置くことは容易ではない。また、データを臨床的に用いるため（患者間を比較する）には、対象ＲＮＡはものさしの役割を果たすことになるため、常に一定の値を示すもので、かつ、安定的に供給できなくてはならない。

そこで、これらの条件を満たすＲＮＡの供給源として、通常は培養細胞を何種類か混合したものを用いるのだが、培養細胞は、もともとの由来する臓器本来のＤＮＡ発現に加え（臓器ごとでハウスキーピング遺伝子の発現レベルが異なるとの報告あり）、さらに細胞株樹立の工程で本来の遺伝子発現とは異なる発現態様を示すことが知られている。そのため一般的なメンテナンス遺伝子（ハウスキーピング遺伝子）では、スケール化（ノーマライズ）が出来ない。

また、ノーマライズ用に、内部標準ではなく、外部標準ＲＮＡ（スパイクＲＮＡ）を検体と対照ＲＮＡに加えるやりかたもある。具体例として、Ｌｕｃｙｄｅａ　Ｍｉｃｒｏａｒｒａｙ　ＳｃｏｒｅＣａｒｄ　ｖ１．１（製品名）（ａｍｅｒｓｈａｍ　ｐｈａｒｍａｃｉａ　ｂｉｏｔｅｃｈ（会社名））においては、ネガティブ・コントロール、ポジティブ・コントロール（ハウスキーピング遺伝子など）以外にダイナミック・レンジ用およびレイシオ・コントロール用に一種類の適切なスパイクＲＮＡを用い、コンピュータ解析と合わせて測定精度の向上が図られている（Ｌｕｃｙｄｅａ　Ｍｉｃｒｏａｒｒａｙ　ＳｃｏｒｅＣａｒｄ　ｖ１．１取扱説明書）。しかし、この方法では、最終的に検体、対照それぞれに加えたスパイクＲＮＡが同等でなければならないが、それを確認する手段がないた

め、あまり推奨されていない（Ｓｔｅｅｎ Ｋｎｕｄｓｅｎ著「ＤＮＡマイクロアレイデータ解析入門」第３章、羊土社 ２００２年）。しかし、スパイクＲＮＡは外部からコントロールできるため、もしうまく補正する手段があれば、非常に有効な対照となり得る。また例えば、“無処理Ａ細胞と薬剤Ｃ処理Ａ細胞”と“無処理Ｂ細胞と薬剤Ｃ処理Ｂ細胞“を実験に使用した場合、各々の”無処理の細胞“をコントロールとして実験を行うと、ＤＮＡマイクロアレイでは薬剤Ｃに対して応答する遺伝子の変動を調べることができるが、細胞Ａと細胞Ｂのもともとの差異は解析しきれない。そこで細胞Ａと細胞Ｂとは、無関係である細胞Ｄを設定して、細胞Ａ、細胞Ｂに対し、細胞Ｄをコントロールとしたデータをそれぞれ取得し、比較することにより、細胞Ａと細胞Ｂの相違が明らかにすることができる。このような場合は細胞Ｄが細胞Ａ、細胞Ｂと無関係なため、一般的なハウスキーピング遺伝子においてもその発現状態は異なる可能性が高く、外部からのスパイクＲＮＡを用いるか、あるいは内部の遺伝子を用いる場合は、あらかじめ最適な補正用遺伝子を選択しておく必要がある。

これまではＤＮＡマイクロアレイはもっぱら研究用に使用されていたため、癌組織と正常組織等、１個体内で同一臓器での比較なので、ハウスキーピング遺伝子での合わせ込み（ノーマライズまたはスケール化）は有効であったが、ＤＮＡマイクロアレイを産業的に利用するためには、アレイの種類・目的によって、まず、使用すべきメンテナンス遺伝子の選択をする必要がある。もし適切な遺伝子を選択できなかった場合は、それらを用いたノーマライズは、逆に誤差の原因となるからである。

また、上記（２）の問題として、特に低発現遺伝子におけるばらつきがある。その原因の一つとして、測定機器（スキャナ装置等）の物理的特性から、測定可能な範囲が非常に限定されていることが挙げられる。

例えば現在市販のマイクロアレイ用スキャナの測定レンジは$10^4$程度で、実際に遺伝子の発現量をマイクロアレイで調べると、高発現から低発現まで同時に同じレーザーパワー（蛍光物質を励起させるためのレーザー光の強さ）、同じＰＭＴゲイン（光電子増倍管の感度：蛍光を測定する感度）では測定できない。例えば、低発現遺伝子を測定するためには、レーザーパワーまたはＰＭＴゲインを上げる必要があり、この時、高発現遺伝子の測定値は飽和しており測定できない。一方、高発現遺伝子を測定するためには、レーザーパワーとＰＭＴゲインを下げる必要性があるが、この時は、低発現遺伝子は測定値がゼロまたはそれに近い値となり、個々のスポットの微妙な差は数値化されない。すなわち、同時に高発現遺伝子と低発現遺伝子を１回のスキャニング条件で測定することはできない。そこで通常のスキャニング方法では発光強度の中間点に測定レンジを調整し、発現が極端に高いものと低いものは、データとして捨てざるを得ないことになる。

さらに、二蛍光標識法を用いても、実際には、スキャナ特性や色素の物理化学的特性、蛍光物質（Ｃｙ３、Ｃｙ５）の核酸への取り込み効率などによってこれらの値にバイアスがかかり、ばらつき以外の誤差の原因となっていることが知られている（例；１．Ｙｕ Ｊ，Ｏｔｈｍａｎ ＭＩ，Ｆａｒｊｏ Ｒ，Ｚａｒｅｐａｒｓｉ Ｓ，ＭａｃＮｅｅ ＳＰ，Ｙｏｓｈｉｄａ Ｓ，Ｓｗａｒｏｏｐ Ａ．Ｅｖａｌｕａｔｉｏｎ ａｎｄ ｏｐｔｉｍｉｚａｔｉｏｎ ｏｆ ｐｒｏｃｅｄｕｒｅｓ ｆｏｒ ｔａｒｇｅｔ ｌａｂｅｌｉｎｇ ａｎｄ ｈｙｂｒｉｄｉｚａｔｉｏｎ ｏｆ ｃＤＮＡ ｍｉｃｒｏａｒｒａｙｓ．Ｍｏｌ Ｖｉｓ．２００２ Ａｐｒ ２６；８：１３０－７．２．’ｔ Ｈｏｅｎ ＰＡ，ｄｅ Ｋｏｒｔ Ｆ，ｖａｎ Ｏｍｍｅｎ ＧＪ，ｄｅｎ Ｄｕｎｎｅｎ ＪＴ．Ｆｌｕｏｒｅｓｃｅｎｔ ｌａｂｅｌｌｉｎｇ ｏｆ ｃＲＮＡ ｆｏｒ ｍｉｃｒｏａｒｒａｙ ａｐｐｌｉｃａｔｉｏｎｓ．Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２００３ Ｍａｒ １；３１（５）：ｅ２０．）。

その結果として低発現および高発現領域のデータがしばしば非線形（非直線性）となり、それら非線形部分のデータが使用できなくなる。そこでそれらの誤差を取り除き、データ全体を線形とするための何らかの工夫が必要となる。Ｓｃｈａｄｔらはこの問題を曲線の部分ごとにスケール化係数を変える局所スケール化をすることで解消できるとしているが（Ｓｃｈａｄｔ，Ｅ．ｅｔ，ａｌ，，２０００，Ａｎａｌｙｚｉｎｇ ｈｉｇｈ－ｄｅｎ

ｓｉｔｙ ｏｌｉｇｏｎｕｃｌｅｏｔｉｄｅ ｇｅｎｅ ｅｚｐｒｅｓｓｉｏｎ ａｒｒａｙ ｄａｔａ，Ｊ．Ｃｅｌｌ．ｂｉｏＣｈｅｍ．，８０：１９２－２０１）、これはＡｆｆｙｍｅｔｒｉｘ（会社名）の合成オリゴタイプのマイクロアレイについてであり、一般的なｃＤＮＡアレイには応用できない。

一方この問題はさらに異なる局面からの困難さがともなっている。例えば上述した「ＤＮＡマイクロアレイデータ解析入門」の日本語版監訳者がその「監訳者の序」において「・・一番大きな問題点は得られた情報の処理、解析、・・・などの過程に研究者がナイーブな点にあると考えられる。・・・その結果の解析（アレイインフォマテイクス）が最も大きな障害となることが予測される。」（羊土社、Ｓｔｅｅｎ Ｋｎｕｄｓｅｎ著、塩島ら監訳、２００２年）と述べているように、ＤＮＡマイクロアレイの解析は、その研究を担う生物学研究者にとって未知の領域であるため、ほとんどの研究者が実験の手技そのものよりも、実験で得られたデータの解析に困難さを感じているのが実情である。

例えばＤＮＡマイクロアレイデータを処理するのに、ほとんどの研究者は既存あるいは市販のソフトウエアパッケージを使用しているが、問題は解析のツール取得ではなく、むしろそこに並べられている様々な解析法のどれを用いて数字化すべきなのか、あるいは数字化したデータをどの方法で解析すべきか（例えば、クラスター解析ひとつにしても階層型クラスター化、ｋ－ｍｅａｎｓクラスター化、自己組織化マップ（ＳＯＭ）等いくつかの方法があり、方法によって結果が異なる等）というところにある。

特にデータの補正方法は、マイクロアレイ全体の補正としてはハウスキーピング遺伝子を用いた補正やグローバル補正等がすでに知られているが（「ＤＮＡマイクロアレイと最新ＰＣＲ法」ｐ３６、細胞工学別冊、秀潤社）、ＤＮＡマイクロアレイに固有の問題、例えばスキャナーの最適領域を外れた値をどのように補正するかについては既存の統計処理方法をそのまま当てはめることができず、それぞれの工夫が必要となる。しかしそれらは一般の生物学研究者にはなじみのない、高度に専門化された数学的手法の理解を必要とするものである（例；Ｍｉｃｒｏａｒｒａｙ ｄａｔａ ｎｏｒｍａｌｉｚａｔｉｏｎ ａｎｄ ｔｒａｎｓｆｏｒｍａｔｉｏｎ，Ｊｏｈｎ Ｑｕａｃｋｅｎｂｕｓｈ，Ｎａｔｕｒｅ Ｇｅｎｅｔｉｃｓ Ｓｕｐｐｌｅｍｅｎｔ，Ｖｏｌ ３２，４９６－５０１，２００２）。

上記のような検討が行われているものの、更なる測定の精度や感度あるいは再現性の向上が求められており、特に遺伝子プロファイリングや病態の遺伝子の解析においては、低発現遺伝子の測定精度や再現性の向上およびそのための簡易で操作性のよいツールは重要な課題となっている。また、同時に作業の煩雑さを軽減し効率化を図る上で、スキャニングからデータ解析に至る完全自動化も重要な課題となっている。

従って、本発明は、上記課題に鑑みてなされたもので、従来は一種類とされていたスパイクＲＮＡやハウスキーピング遺伝子などのコントロール用遺伝子に複数の遺伝子ＲＮＡを使用することにより、ＤＮＡマイクロアレイによる測定精度や再現性の向上を可能とし、測定レンジを拡大させ、

また、ＲＮＡの増幅や標識反応、アレイ上での結合（ハイブリダイゼーション）反応、およびシグナル取込み等の一連の過程において現れる様々なばらつきや誤差を、補正によって取り除くことにより信頼性の高いデータを取得し、

さらにそれらの補正方法を一般生物学研究者が容易に理解しうる構成とすることで、ユーザーが自由に操作・活用でき、

同時に作業の煩雑さを軽減し効率化を図り、スキャニングからデータ解析に至る完全自動化を可能とするＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体を提供することを目的としている。

発明の開示

すなわち、本発明は従来一種類とされていたスパイクＲＮＡやハウスキーピング遺伝子などのコントロール用遺伝子に複数の遺伝子ＲＮＡを使用することにより、より具体的には一例として、ウミシイタケルシフェラーゼ（Ｒｅｎｉｌｌａｌｕｃｉｆｅｒａｓｅ）遺伝

子（Ｒｌｕｃ）、バキュロウィルス（ｂａｃｕｌｏｖｉｒｕｓ）膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の三種の遺伝子断片由来ｍＲＮＡをスパイクＲＮＡとして用いること、または、ｂｅｔａ－ａｃｔｉｎ、ＧＡＰＤＨ、ｐｏｌｙ（Ａ） ｂｉｎｄｉｎｇ ｐｒｏｔｅｉｎ（ＰＡＢＰ）、ｇｌｙｃｙｌ ｔＲＮＡ ｓｙｎｔｈｅｔａｓｅ、ｈｎＲＮＰ Ｈ１等の遺伝子をコントロール用スポットとして用いることにより、本発明を完成するに至った。

また、本発明は既存の一律的な統計処理方法ではなく、実際に即した一連のデータ補正・処理方法を統合して用いることにより、本発明を完成するに至った。

上述した目的を達成するため、本発明にかかるＤＮＡマイクロアレイデータ処理方法は、複数のスパイク用遺伝子から調整したＲＮＡまたは複数の人工的に合成したＲＮＡをコントロールとして測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度に基づいてスキャニング画像の上記測定対象のＲＮＡの各スポットに対してシグナル強度のノーマライゼーションを行うことを特徴とする。

この方法によれば、複数のスパイク用遺伝子から調整したＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度に基づいてスキャニング画像の測定対象のＲＮＡの各スポットに対してシグナル強度のノーマライゼーションを行うので、ＤＮＡマイクロアレイによる測定レンジの拡大、精度や再現性の向上が図られ、特に遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。

また、これにより、同作業の煩雑さを軽減し効率化を図ることができ、スキャニングからデータ解析に至るまでの完全自動化をも図ることができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、測定対象のＲＮＡに複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロール用ＲＮＡを混入して調整した試料に対してＤＮＡマイクロアレイを用いてハイブリダイゼーションを行うハイブリダイゼーションステップと、上記ハイブリダイゼージョンステップにてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングするスキャニングステップと、上記スキャニングステップにてスキャニングされたスキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズステップと、上記ノーマライズステップにて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップとを含むことを特徴とする。

この方法によれば、測定対象のＲＮＡに複数のコントロール用遺伝子から調整したＲＮＡを混入して調整した試料に対してＤＮＡマイクロアレイを用いてハイブリダイゼーションを行い、ハイブリダイゼーションを行ったＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングし、スキャニングされたスキャニング画像の複数のコントロール用ＲＮＡが結合した各スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションをコントロール用ＲＮＡ毎に実行し、各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、ＤＮＡマイクロアレイによる測定レンジの拡大、精度や再現性の向上が図られ、特に遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることがで

きる。
また、これにより、同作業の煩雑さを軽減し効率化を図ることができ、スキャニングからデータ解析に至るまでの完全自動化をも図ることができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記スキャニングステップは、上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用ＲＮＡ毎に上記発光支持体をスキャニングすることを特徴とする。
これはスキャニングの一例を一層具体的に示すものである。この方法によれば、複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用ＲＮＡ毎に発光支持体をスキャニングするので、それぞれのコントロール用ＲＮＡが一番安定したシグナルとして取り込めるような条件で、コントロール毎に感度調整を行いながら複数回（用いたコントロールの数）スキャニングを行うことができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。
これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。この方法によれば、測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、非特異的な反応を抑え、ヒトの遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。より具体的には、例えばヒトのＣ型肝炎のインターフェロンによる治療効果の予測判定などにＤＮＡマイクロアレイが利用可能となり、従来から用いられている肝生検などで得られている病理診断情報に加え、積極的な治療を目指した患者への情報提供を提供することができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合で、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。
これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。この方法によれば、測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合で、複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、さらに測定対象を非ヒト遺伝子まで広げることにより、例えばヒト組織中のウイルスの存在とその量を確認しつつヒト遺伝子の発現をモニターすることなどが可能となる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いることを特徴とする。
これはコントロール用ＲＮＡの一例を一層具体的に示すものである。この方法によれば、複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整することを特徴とする。

これはコントロール用ＲＮＡの一例を一層具体的に示すものである。この方法によれば、複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）を用いて調整するので、これらをスパイクとして高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群について適切な測定データを取得することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のコントロール用スポットに結合したＲＮＡのシグナル強度に基づいてスキャニング画像の測定対象のＲＮＡが結合した各スポットに対してシグナル強度のノーマライゼーションを行うことを特徴とする。

この方法によれば、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、当該複数のコントロール用スポットに結合したＲＮＡのシグナル強度に基づいてスキャニング画像の測定対象のＲＮＡが結合した各スポットに対してシグナル強度のノーマライゼーションを行うので、コントロール用遺伝子は必ずしも外部から加えるものに限られず、測定対象サンプル内に含まれ、かつサンプル毎の変動がほとんど見られない遺伝子（例えばいわゆるハウスキーピング遺伝子類等）を選択し、コントロールとして同様に用いることができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、測定対象遺伝子のＲＮＡに対して、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行うハイブリダイゼーションステップと、上記ハイブリダイゼージョンステップにてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングするスキャニングステップと、上記スキャニングステップにてスキャニングされたスキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズステップと、上記ノーマライズステップにて各コントロール用スポットによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップと含むことを特徴とする。

この方法によれば、測定対象遺伝子のＲＮＡに対して、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行い、ハイブリダイゼーションを行ったＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングし、スキャニングされたスキャニング画像の複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションをコントロール用スポットの遺伝子毎に実行し、各コントロール用スポットによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測

定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、コントロール用遺伝子は必ずしも外部から加えるものに限られず、測定対象サンプル内に含まれ、かつサンプル毎の変動がほとんど見られない遺伝子（例えばいわゆるハウスキーピング遺伝子類等）を選択し、コントロールとして同様に用いることができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記スキャニングステップは、上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎に上記発光支持体をスキャニングすることを特徴とする。

これはスキャニングの一例を一層具体的に示すものである。この方法によれば、複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用スポットの遺伝子毎に発光支持体をスキャニングするので、それぞれのコントロール用スポットが一番安定したシグナルとして取り込めるような条件で、コントロール用スポット毎に感度調整を行いながら複数回（用いたコントロール用スポットの数）スキャニングを行うことができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いることを特徴とする。

これはコントロール用スポットの一例を一層具体的に示すものである。この方法によれば、複数のコントロール用ＲＮＡは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記ノーマライズステップは、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化ステップをさらに含むことを特徴とする。

これはノーマライズステップの一例を一層具体的に示すものである。この方法によれば、複数の異なるコントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類のコントロール用スポットのシグナル強度に換算することにより統合するので、複数の異なる条件で取得した各スポットのシグナル強度を容易に比較することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記ノーマライズステップは、上記統合化ステップにて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定ステップと、上記棄却値選定ステップにより選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定ステップと、上記仮代表値決定ステップにて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定する回帰式補正ステップとをさらに含むことを特徴とする。

これはノーマライズステップの一例を一層具体的に示すものである。この方法によれば、換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基

づき棄却値を選定し、選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定し、決定された仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定するので、従来技術における測定上の誤差の問題、特に低発現領域と高発現領域で見られる非線形の問題を解消することができる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理方法は、上記に記載のＤＮＡマイクロアレイデータ処理方法において、上記ノーマライズステップは、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正ステップをさらに含むことを特徴とする。

これはノーマライズステップの一例を一層具体的に示すものである。この方法によれば、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、対照検体と測定検体の測定上の誤差を補正して代表値を決定するので、さらに測定誤差を少なくすることができる。

また、本発明はＤＮＡマイクロアレイデータ処理システムに関するものであり、本発明にかかるＤＮＡマイクロアレイデータ処理システムは、ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置と、当該スキャナ装置によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置とを含んで構成されるＤＮＡマイクロアレイデータ処理システムであって、上記スキャナ装置は、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または人工的に合成したコントロール用ＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャニング手段を備え、上記ＤＮＡマイクロアレイデータ処理装置は、上記スキャニング手段にてスキャニングされたスキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズ手段と、上記ノーマライズ手段にて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段とを備えることを特徴とする。

このシステムによれば、ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置と、当該スキャナ装置によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置とを含んで構成されるＤＮＡマイクロアレイデータ処理システムであって、スキャナ装置は、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または人工的に合成したコントロール用ＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、ＤＮＡマイクロアレイデータ処理装置は、スキャニングされたスキャニング画像の複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションをコントロール用ＲＮＡ毎に実行し、各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、ＤＮＡマイクロアレイによる測定レンジの拡大、精度や再現性の向上が図られ、特に遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。

また、これにより、同作業の煩雑さを軽減し効率化を図ることができ、スキャニングからデータ解析に至るまでの完全自動化をも図ることができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマ

イクロアレイデータ処理システムにおいて、上記スキャニング手段は、上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように感度を調整して上記コントロール用ＲＮＡ毎に上記発光支持体をスキャニングするスキャニング感度調整手段をさらに備えることを特徴とする。
これはスキャニングの一例を一層具体的に示すものである。このシステムによれば、複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用ＲＮＡ毎に上記発光支持体をスキャニングするので、それぞれのコントロール用ＲＮＡが一番安定したシグナルとして取り込めるような条件で、コントロール毎に感度調整を行いながら複数回（用いたコントロールの数）スキャニングを行うことができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡから調整したＲＮＡであることを特徴とする。
これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。このシステムによれば、測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、非特異的な反応を抑え、ヒトの遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。より具体的には、例えばヒトのＣ型肝炎のインターフェロンによる治療効果の予測判定などにＤＮＡマイクロアレイが利用可能となり、従来から用いられている肝生検などで得られている病理診断情報に加え、積極的な治療を目指した患者への情報提供を提供することができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。
これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。このシステムによれば、測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、さらに測定対象を非ヒト遺伝子まで広げることにより、例えばヒト組織中のウイルスの存在とその量を確認しつつヒト遺伝子の発現をモニターすることなどが可能となる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いることを特徴とする。
これはコントロール用ＲＮＡの一例を一層具体的に示すものである。このシステムによれば、複数のコントロール用ＲＮＡは、測定対象の遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いるので、非特異的反応を抑え、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整することを特

徴とする。
これはコントロール用ＲＮＡの一例を一層具体的に示すものである。このシステムによれば、複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整するので、これらをスパイクとして高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群について適切な測定データを取得することができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置と、当該スキャナ装置によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置とを含んで構成されるＤＮＡマイクロアレイデータ処理システムであって、上記スキャナ装置は、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャニング手段を備え、上記ＤＮＡマイクロアレイデータ処理装置は、上記スキャニング手段にてスキャニングされたスキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポット遺伝子毎に実行するノーマライズ手段と、上記ノーマライズ手段にて各コントロール用スポットによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段とを備えることを特徴とする。
このシステムによれば、ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置と、当該スキャナ装置によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置とを含んで構成されるＤＮＡマイクロアレイデータ処理システムであって、スキャナ装置は、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてそれぞれスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行い、ＤＮＡマイクロアレイデータ処理装置は、スキャニングされたスキャニング画像の複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポット遺伝子毎に実行し、各コントロール用スポットによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、コントロール用遺伝子は必ずしも外部から加えるものに限られず、測定対象サンプル内に含まれ、かつサンプル毎の変動がほとんど見られない遺伝子（例えばいわゆるハウスキーピング遺伝子類等）を選択し、コントロールとして同様に用いることができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記スキャニング手段は、上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎に上記発光支持体をスキャニングすることを特徴とする。
これはスキャニングの一例を一層具体的に示すものである。このシステムによれば、複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用スポットの遺伝子毎に発光支持体をスキャニングするので、それぞれのコントロール用スポットが一番安定したシグナルとして取り込めるような条件で、コントロール用スポット毎に感度調整を行いながら複数回（用いたコントロール用スポットの数）スキャニングを行

うことができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いることを特徴とする。
これはコントロール用スポットの一例を一層具体的に示すものである。このシステムによれば、複数のコントロール用ＲＮＡは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記ノーマライズ手段は、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化手段をさらに含むことを特徴とする。
これはノーマライズ手段の一例を一層具体的に示すものである。このシステムによれば、複数の異なるコントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類のコントロール用スポットのシグナル強度に換算することにより統合するので、複数の異なる条件で取得した各スポットのシグナル強度を容易に比較することができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記ノーマライズ手段は、上記統合化手段にて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定手段と、上記棄却値選定手段により選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定手段と、上記仮代表値決定手段にて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定する回帰式補正手段とをさらに含むことを特徴とする。
これはノーマライズ手段の一例を一層具体的に示すものである。このシステムによれば、換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定し、選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定し、決定された仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定するので、従来技術における測定上の誤差の問題、特に低発現領域と高発現領域で見られる非線形の問題を解消することができる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理システムは、上記に記載のＤＮＡマイクロアレイデータ処理システムにおいて、上記ノーマライズ手段は、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正手段をさらに含むことを特徴とする。
これはノーマライズ手段の一例を一層具体的に示すものである。このシステムによれば、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、対照検体と測定検体の測定上の誤差を補正して代表値を決定するので、さらに測定誤差を少なくすること

ができる。
また、本発明はＤＮＡマイクロアレイデータ処理装置に関するものであり、本発明にかかるＤＮＡマイクロアレイデータ処理装置は、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロールＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得手段と、上記スキャニング画像取得手段にて取得された上記スキャニング画像の上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズ手段と、上記ノーマライズ手段にて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段とを備えることを特徴とする。
この装置によれば、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロールＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度に基づいてスキャニング画像の測定対象のＲＮＡの各スポットに対してシグナル強度のノーマライゼーションを行うので、ＤＮＡマイクロアレイによる測定レンジの拡大、精度や再現性の向上が図られ、特に遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。さらに測定対象を非ヒト遺伝子まで広げることにより、例えばヒト組織中のウイルスの存在とその量を確認しつつヒト遺伝子の発現をモニターすることなどが可能となる。
また、これにより、同作業の煩雑さを軽減し効率化を図ることができ、スキャニングからデータ解析に至るまでの完全自動化をも図ることができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記スキャニング画像は、上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用ＲＮＡ毎にスキャニングされた複数の画像であることを特徴とする。
これはスキャニングの一例を一層具体的に示すものである。この装置によれば、複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用ＲＮＡ毎に上記発光支持体をスキャニングするので、それぞれのコントロール用ＲＮＡが一番安定したシグナルとして取り込めるような条件で、コントロール毎に感度調整を行いながら複数回（用いたコントロールの数）スキャニングを行うことができるようになる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。
これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。この装置によれば、測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、非特異的な反応を抑え、ヒトの遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。より具体的には、例えばヒトのＣ型肝炎のインターフェロンによる治療効果の予測判定などにＤＮＡマイクロアレイが利用可能となり、

従来から用いられている肝生検などで得られている病理診断情報に加え、積極的な治療を目指した患者への情報提供を提供することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。

これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。この装置によれば、測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、さらに測定対象を非ヒト遺伝子まで広げることにより、例えばヒト組織中のウイルスの存在とその量を確認しつつヒト遺伝子の発現をモニターすることなどが可能となる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記複数のコントロール用ＲＮＡは、測定対象遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いることを特徴とする。

これはコントロール用ＲＮＡの一例を一層具体的に示すものである。この装置によれば、複数のコントロール用ＲＮＡは、測定対象遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整することを特徴とする。

これはコントロール用ＲＮＡの一例を一層具体的に示すものである。この装置によれば、複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整するので、これらをスパイクとして高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群について適切な測定データを取得することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得手段と、上記スキャニング画像取得手段にて取得された上記スキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズ手段と、上記ノーマライズ手段にて各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段とを備えることを特徴とする。

この装置によれば、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得し、取得されたスキャニング画像の

複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションをコントロール用スポットの遺伝子毎に実行し、各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、コントロール用遺伝子は必ずしも外部から加えるものに限られず、測定対象サンプル内に含まれ、かつサンプル毎の変動がほとんど見られない遺伝子（例えばいわゆるハウスキーピング遺伝子類等）を選択し、コントロールとして同様に用いることができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記スキャニング画像は、上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎にスキャニングされた複数の画像であることを特徴とする。

これはスキャニング画像の一例を一層具体的に示すものである。この装置によれば、スキャニング画像は、複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用スポットの遺伝子毎にスキャニングされた複数の画像であるので、それぞれのコントロール用スポットが一番安定したシグナルとして取り込めるような条件で、コントロール用スポット毎に感度調整を行いながら複数回（用いたコントロール用スポットの数）スキャニングを行うことができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いることを特徴とする。

これはコントロール用スポットの一例を一層具体的に示すものである。この装置によれば、複数のコントロール用ＲＮＡは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記ノーマライズ手段は、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化手段をさらに含むことを特徴とする。

これはノーマライズ手段の一例を一層具体的に示すものである。この装置によれば、複数の異なるコントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類のコントロール用スポットのシグナル強度に換算することにより統合するので、複数の異なる条件で取得した各スポットのシグナル強度を容易に比較することができるようになる。

つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記ノーマライズ手段は、上記統合化手段にて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定手段と、上記棄却値選定手段により選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定手段と、上記仮代表値決定手段にて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して

代表値を決定する回帰式補正手段とをさらに含むことを特徴とする。
これはノーマライズ手段の一例を一層具体的に示すものである。この装置によれば、換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定し、選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定し、決定された仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定するので、従来技術における測定上の誤差の問題、特に低発現領域と高発現領域で見られる非線形の問題を解消することができる。
つぎの発明にかかるＤＮＡマイクロアレイデータ処理装置は、上記に記載のＤＮＡマイクロアレイデータ処理装置において、上記ノーマライズ手段は、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正手段をさらに含むことを特徴とする。
これはノーマライズ手段の一例を一層具体的に示すものである。この装置によれば、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、対照検体と測定検体の測定上の誤差を補正して代表値を決定するので、さらに測定誤差を少なくすることができる。
また、本発明はプログラムに関するものであり、本発明にかかるプログラムは、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロールＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得ステップと、上記スキャニング画像取得ステップにて取得された上記スキャニング画像の上記複数のコントロール用ＲＮＡが結合した各スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用ＲＮＡ毎に実行するノーマライズステップと、上記ノーマライズステップにて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップとを含むＤＮＡマイクロアレイデータ処理方法をコンピュータに実行させることを特徴とする。
このプログラムによれば、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロールＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行い、複数のコントロール用ＲＮＡが結合した各スポットのシグナル強度に基づいてスキャニング画像の測定対象のＲＮＡの各スポットに対してシグナル強度のノーマライゼーションを行い、各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、ＤＮＡマイクロアレイによる測定レンジの拡大、精度や再現性の向上が図られ、特に遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。さらに測定対象を非ヒト遺伝子まで広げることにより、例えばヒト組織中のウイルスの存在とその量を確認しつつヒト遺伝子の発現をモニターすることなどが可能となる。
また、これにより、同作業の煩雑さを軽減し効率化を図ることができ、スキャニングからデータ解析に至るまでの完全自動化をも図ることができるようになる。
つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記スキャニング画像は、上記複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用ＲＮＡ毎にスキャニングされた複数の画像であることを特徴とする。

これはスキャニングの一例を一層具体的に示すものである。このプログラムによれば、複数のコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用ＲＮＡ毎に上記発光支持体をスキャニングするので、それぞれのコントロール用ＲＮＡが一番安定したシグナルとして取り込めるような条件で、コントロール毎に感度調整を行いながら複数回（用いたコントロールの数）スキャニングを行うことができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、上記複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。

これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。このプログラムによれば、測定対象のＲＮＡはヒト遺伝子ＲＮＡであり、複数のコントロール用ＲＮＡはヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、非特異的な反応を抑え、ヒトの遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。より具体的には、例えばヒトのＣ型肝炎のインターフェロンによる治療効果の予測判定などにＤＮＡマイクロアレイが利用可能となり、従来から用いられている肝生検などで得られている病理診断情報に加え、積極的な治療を目指した患者への情報提供を提供することができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、上記複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであることを特徴とする。

これは測定対象のＲＮＡおよびコントロール用ＲＮＡの一例を一層具体的に示すものである。このプログラムによれば、測定対象のＲＮＡは非ヒト遺伝子ＲＮＡ、または非ヒト遺伝子ＲＮＡとヒト遺伝子ＲＮＡの混合であり、複数のコントロール用ＲＮＡは当該非ヒト遺伝子またはヒト遺伝子のいずれともハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡおよび／または複数の人工的に合成したＲＮＡであるので、さらに測定対象を非ヒト遺伝子まで広げることにより、例えばヒト組織中のウイルスの存在とその量を確認しつつヒト遺伝子の発現をモニターすることなどが可能となる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記複数のコントロール用ＲＮＡは、測定対象遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いることを特徴とする。

これはコントロール用ＲＮＡの一例を一層具体的に示すものである。このプログラムによれば、複数のコントロール用ＲＮＡは、測定対象遺伝子群の高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する量のＲＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整することを特徴とする。

これはコントロール用ＲＮＡの一例を一層具体的に示すものである。このプログラムによれば、複数のコントロール用ＲＮＡは、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片を用いて調整するので、これらをスパイクとして高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群について適切な測定データを取得することができるようになる。

つぎの発明にかかるプログラムは、測定対象サンプルに含まれ、かつ測定対象遺伝子以外

から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行った上記ＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得ステップと、上記スキャニング画像取得ステップにて取得された上記スキャニング画像の上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、上記測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズステップと、上記ノーマライズステップにて各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択ステップとを含むことを特徴とする。

このプログラムによれば、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得し、取得されたスキャニング画像の複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションをコントロール用スポットの遺伝子毎に実行し、各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するので、コントロール用遺伝子は必ずしも外部から加えるものに限られず、測定対象サンプル内に含まれ、かつサンプル毎の変動がほとんど見られない遺伝子（例えばいわゆるハウスキーピング遺伝子類等）を選択し、コントロールとして同様に用いることができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記スキャニング画像は、上記複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるように調整して上記コントロール用スポットの遺伝子毎にスキャニングされた複数の画像であることを特徴とする。

これはスキャニング画像の一例を一層具体的に示すものである。このプログラムによれば、スキャニング画像は、複数のコントロール用スポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度（例えば、スキャナ装置の上限値の半分程度）になるように調整してコントロール用スポットの遺伝子毎にスキャニングされた複数の画像であるので、それぞれのコントロール用スポットが一番安定したシグナルとして取り込めるような条件で、コントロール用スポット毎に感度調整を行いながら複数回（用いたコントロール用スポットの数）スキャニングを行うことができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いることを特徴とする。

これはコントロール用スポットの一例を一層具体的に示すものである。このプログラムによれば、複数のコントロール用ＲＮＡは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いるので、各遺伝子の発現レベルに従って適切な測定データを取得することができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記ノーマライズステップは、上記複数の異なる上記コントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類の上記コントロール用スポットのシグナル強度に換算することにより統合する統合化ステップをさらに含むことを特徴とする。

これはノーマライズステップの一例を一層具体的に示すものである。このプログラムによ

れば、複数の異なるコントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なる上記コントロール用ＲＮＡの中から選択した一種類または数種類のコントロール用スポットのシグナル強度に換算することにより統合するので、複数の異なる条件で取得した各スポットのシグナル強度を容易に比較することができるようになる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記ノーマライズステップは、上記統合化ステップにて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定ステップと、上記棄却値選定ステップにより選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定する仮代表値決定ステップと、上記仮代表値決定ステップにて決定された上記仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定する回帰式補正ステップとをさらに含むことを特徴とする。

これはノーマライズステップの一例を一層具体的に示すものである。このプログラムによれば、換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定し、選定された上記棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値を得て各遺伝子の仮代表値を決定し、決定された仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正して代表値を決定するので、従来技術における測定上の誤差の問題、特に低発現領域と高発現領域で見られる非線形の問題を解消することができる。

つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記ノーマライズステップは、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、上記対照検体と上記測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正ステップをさらに含むことを特徴とする。

これはノーマライズステップの一例を一層具体的に示すものである。このプログラムによれば、あらかじめ選択した、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を利用して回帰式を求めて全ての仮代表値を補正することにより、対照検体と測定検体の測定上の誤差を補正して代表値を決定するので、さらに測定誤差を少なくすることができる。

また、本発明は記録媒体に関するものであり、本発明にかかる記録媒体は、上記に記載されたプログラムを記録したことを特徴とする。

この記録媒体によれば、当該記録媒体に記録されたプログラムをコンピュータに読み取らせて実行することによって、上記に記載されたプログラムをコンピュータを利用して実現することができ、これら各プログラムと同様の効果を得ることができる。

発明を実施するための最良の形態

以下に、本発明にかかるＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体の実施の形態を図面に基づいて詳細に説明する。なお、この実施の形態によりこの発明が限定されるものではない。また、以下の実施の形態においては、コントロール用遺伝子としてスパイク用遺伝子を用いる場合やハウスキーピング遺伝子等をコントロールスポットに用いる場合を一例に説明するが、本発明はかかる場合に限定されるものではない。

［本発明の概要］

以下、本発明の概要について説明し、その後、本発明の構成および処理等について図１から図４を用いて詳細に説明する。図１は本発明の基本原理を示す原理構成図である。

本発明は、概略的に、以下の基本的特徴を有する。すなわち、本発明は、外部コントロールとして従来一種類とされていたスパイクＲＮＡに複数のＲＮＡ（例えば、高・中・低発現群に相当する３種類の人工ＲＮＡであり、より具体的には一例として、ウミシイタケル

シフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片）を用いてＤＮＡマイクロアレイによるハイブリダイゼーションを行う（ステップＳＡ－１）。

人工ＲＮＡは、昆虫、細菌、ウイルスなど、目的とするマイクロアレイと相同性が低くクロスハイブリダイゼーションが起こらない遺伝子配列であれば利用可能である。相同性及びクロスハイブリダイゼーションについてはコンピューター上である程度予測することができる。

図２は、複数のスパイクＲＮＡを用いた二蛍光標識法によるスパイク実験の概念を説明する概念図である。図２に示すように、まず、測定対象のＲＮＡ（例えば、ウイルス感染ヒト肝細胞から抽出および増幅したＲＮＡ等）とコントロールＲＮＡ（例えば、正常ヒト肝細胞から抽出および増幅したＲＮＡ等）に、それぞれコントロール（スパイク）として一定量のスパイク用遺伝子から調整したＲＮＡ（以下、「スパイクＲＮＡ」という。）を加え、逆転写酵素を用いてＣｙ３とＣｙ５でそれぞれ（例えば測定対象ＲＮＡをＣｙ３で、コントロールＲＮＡをＣｙ５で）標識し試料を調整する。標識は必ずしもＣｙ３とＣｙ５に限定するものではなく、ビオチン標識やアミノアリル標識後、他の蛍光物質等を付加させることも可能である。

Ｃｙ３とＣｙ５で標識されたＲＮＡは、混合してＤＮＡマイクロアレイ上にのせるので、各スポットに対し競合的に反応（ハイブリダイズ）することになる。これをＣｙ３またはＣｙ５のそれぞれの波長で発光強度を測定し、スポット毎にＣｙ３とＣｙ５の測定値の比を得ることで、スポット差（例えば各スポットのＤＮＡ量）に関係なく、コントロールＲＮＡに対する測定対象のＲＮＡの増減が把握できる。

スパイクＲＮＡは、例えば、高・中・低発現群に相当する３種類の人工ＲＮＡであり、例えば測定対象のＲＮＡがヒト遺伝子ＲＮＡである場合には、ヒト遺伝子とはハイブリダイズしない複数の非ヒト遺伝子断片から調整したＲＮＡを用いてもよく、より具体的には、一例として、ウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ）、バキュロウィルス膜蛋白遺伝子（ｇｐ６４）、および、ラムダファージｅａ２２遺伝子（ｅａ２２）の断片ＲＮＡを用いてもよい。またコントロール用ＲＮＡは、複数の人工的に合成したＲＮＡでもよい。

ついで、両ターゲット溶液を混合し、ＤＮＡマイクロアレイ上でハイブリダイゼーションを行い洗浄を行う。ここで、ＤＮＡマイクロアレイのスポットとして、各スパイクＲＮＡの専用のスポット（各スパイクにつき複数のスポットを作製してもよい）と、測定対象の遺伝子ＲＮＡ用のスポット（各遺伝子につき複数（例えば、３つ）のスポットを作製してもよい）を作製している。

また、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ってもよい。すなわち、測定対象サンプル内に含まれ、かつサンプル毎の変動がほとんど見られない遺伝子（例えばいわゆるハウスキーピング遺伝子類等）を選択し、コントロールスポットとして作製してもよい。

ここで、複数のコントロール用スポットは、測定対象サンプル内で、それぞれ高発現遺伝子群、中発現遺伝子群、および、低発現遺伝子群に相当する発現量を示す遺伝子から選択された遺伝子のＤＮＡを用いたものでもよく、具体的には例えばｂｅｔａ－ａｃｔｉｎ、ＧＡＰＤＨ、ｐｏｌｙ（Ａ） ｂｉｎｄｉｎｇ ｐｒｏｔｅｉｎ（ＰＡＢＰ）、ｇｌｙｃｙｌ ｔＲＮＡ ｓｙｎｔｈｅｔａｓｅ、ｈｎＲＮＰ Ｈ１等の遺伝子を用いてもよい。

再び、図１に戻り、本発明は、スキャナ装置を用いてＤＮＡマイクロアレイの発光支持体のスキャニングを行い、スパイクＲＮＡ等の各コントロール用ＲＮＡが結合したスポットのシグナル強度や、ハウスキーピング遺伝子等のコントロール用スポットのシグナル強度に基づいてノーマライズ（正規化）処理を実行する（ステップＳＡ－２）。ここで、図３は、スキャナ装置を用いてＤＮＡマイクロアレイの発光支持体のスキャニングを行い、各スパイクＲＮＡのシグナル強度に基づいてノーマライゼーションを行う概念を説明するフロー図である。

図３に示すように、まず、本発明は、ハイブリダイゼーションを行ったＤＮＡマイクロアレイについてスキャナ装置を用いて発光支持体をスキャニングする（ステップＳＢ－１）。

ここで、図３のステップＳＢ－１に示すように、複数のスパイクＲＮＡ等のコントロール用スポットのシグナル強度や、試験管内対象遺伝子等のコントロールスポットに結合したＲＮＡのシグナル強度が予め定めた測定シグナル強度になるようにスキャナ装置の感度を調整してスパイクＲＮＡ毎に発光支持体をスキャニングしてもよい。すなわち、共焦点顕微鏡型スキャナ装置等を用いてスキャニングを行う際に、それぞれのスパイクＲＮＡ等が一番安定したシグナルとして取り込めるような条件で、スパイク等毎にスキャナ装置の感度調整を行いながら複数回（用いたスパイク等の数）スキャニングを行う。なお、図３に示す例では、３種類のスパイクＲＮＡを用いているので、３つのスキャニング画像が作成されている。また、同様に試験管内対象遺伝子等のコントロールスポット毎にスキャニング画像を作成する。

このように適切な感度になるようにスキャナ装置側で感度調整を行いながら各スパイクシグナルを測定することにより、レンジ幅があまり広くないスキャナ装置を用いる場合であってもハードウェア上の検出限界による精度低下等を招くことなくシグナル強度を測定することが可能になり、特に遺伝子プロファイリングや病態遺伝子の解析において重要な低発現遺伝子の測定精度や再現性の向上を図ることができる。

すなわち、スパイクＲＮＡ等は、発現量を変えて複数種類加えてあるので、これまでの方法では発現量が極端に高かったりあるいは低すぎたりして、コントロールＲＮＡ（上記の例ではＣｙ５で標識されている）と測定対象ＲＮＡ（上記の例ではＣｙ３で標識されている）の比がとれない場合でも、それぞれ適切なスパイクを選択することでスキャナの感度調整を行い、より精度の高い測定が、あるいは測定そのものが、可能となる。

ついで、本発明は、コンピュータなどの情報処理装置（ＤＮＡマイクロアレイデータ処理装置）を用いて、スキャナ装置からスキャニング画像を取り込んで画像解析処理等を行う。

まず、本発明は、スキャニング画像について画像解析処理等を施し、各スポットのシグナル強度を算出する（ステップＳＢ－２）。ここで、シグナル強度の算出は、個々のスポットを円で囲んでシグナル強度を算出するｃｉｒｃｌｅ法、スポットの周囲を矩形で囲むｒｅｃｔａｎｇｌｅ法、シグナル強度に応じて自動認識を手法など、いずれの算出手法を用いてもよい。

ついで、本発明は、このように算出したシグナル強度についてバックグラウンド補正を行う（ステップＳＢ－３）。ここで、バックグラウンド補正手法は、各スポットのシグナル強度から各スポットの周囲のブランク領域のバックグラウンド値を引くことによってバックグラウンドの補正を行なう方法（上述のｃｉｒｃｌｅ法等において用いられる）、個々のシグナル強度からブランク領域の平均バックグラウンド値を引くことによってバックグラウンドの補正を行なう方法、上述したｒｅｃｔａｎｇｌｅ法において矩形内の全シグナルのうち下位１０～２５％等のシグナル強度を平均化したものをバックグラウンドとして補正する方法など、いずれの補正手法を用いてもよい。

ついで、本発明は、解析対象となるスポットの抽出を行う（ステップＳＢ－４）。ここで、抽出条件は、例えば、抽出するスポットのシグナル強度の上限値としてスキャナ装置のハードウェア限界値を設定してもよく、下限値として全スポットの平均値に対するχ（$\chi^2$ 統計量の平方根）％のシグナル強度値を設定してもよい。また、ネガティブ・コントロールを用いた場合には、ネガティブ・コントロールのシグナル強度のＳＤ値（標準偏差値）等を用いて下限値を求めてもよい。このように、本発明による解析対象スポットの抽出条件は、各ＤＮＡマイクロアレイの実験ごとの測定データのばらつきや偏りを各種の既知の統計的手法により適宜判定し、ケースバイケースで任意に抽出条件を設定することができる。

ついで、本発明は、スキャニング画像の複数のスパイクＲＮＡ等のシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡの各スポットのシグナル

強度のノーマライゼーションをスパイクＲＮＡ毎に実行する（ステップＳＢ－５）。
ここで、本発明によるノーマライゼーションは以下の手法のうちいずれを用いてもよい。
１）Ｃｙ３側、Ｃｙ５側のスパイクＲＮＡ等のシグナル強度に係数を掛けて、双方のシグナル強度が予め定めた基準シグナル強度（例えば、上限値の２分の１等）になるようにノーマライゼーションを行う。
２）Ｃｙ３側、Ｃｙ５側のスパイクＲＮＡ等のシグナル強度のうちどちらが基準シグナル強度（例えば、上限値の２分の１等）に近いかを判定し、遠い方のシグナル強度を近い方のシグナル強度に合わせるようにノーマライゼーションを行う。
３）Ｃｙ３側、Ｃｙ５側のスパイクＲＮＡ等のシグナル強度のうちどちらかのシグナル強度に一方のシグナル強度を合わせるようにノーマライゼーションを行う。
ついで、本発明は、ノーマライズ後の各測定データを保存する（ステップＳＢ－６）。また、同様に、図３で説明した手法を用いて、スキャニング画像の複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションをコントロール用スポットの遺伝子毎に実行する。
再び、図１に戻り、本発明は、ノーマライズ後の各測定データについてデータマージ処理を行い、各測定対象のＲＮＡのうち各スパイクＲＮＡ等のいずれかに基づいてノーマライズした測定データから最も適切なシグナル強度を抽出する（ステップＳＡ－３）。すなわち、図４を用いて後述するように、各スパイクＲＮＡ等によってノーマライズされた各スポットのシグナル強度に基づいて、スキャニング画像の各測定対象の遺伝子ＲＮＡのシグナル強度のうち最適なシグナル強度を選択する。
ここで、図４は、各測定対象の遺伝子ＲＮＡについて最適なシグナル強度を選択する場合の概念を説明するフロー図である。図４に示すように、まず、本発明は、上述した各スパイク等毎にノーマライズを行った後の測定データを読み込む（ステップＳＣ－１）。
そして、ＤＮＡマイクロアレイ上にスポットされた測定用の遺伝子ＲＮＡの番号を示す添え字であるｉを初期化し（ステップＳＣ－２）、測定データ上のｉ番目の遺伝子ＲＮＡについて全てのスポット（本例では３つのスポット）がＣｙ３側およびＣｙ５側の双方の解析対象上限値および下限値の範囲内に入っているか判定する（ステップＳＣ－３）。
ステップＳＣ－３において入っていると判定された場合には、他のスパイクＲＮＡに基づいてノーマライズされた測定データにおいても入っているかを判定する（ステップＳＣ－４）。
ステップＳＣ－４において入っていると判定された場合には、予め定めた規定値（例えば、上限値の２分の１の値等）に近いシグナル強度となる方の測定データを選択し（ステップＳＣ－５）、ｉ番目の遺伝子のシグナル強度として採用する（ステップＳＣ－６）。
そして、ｉを１加算し（ステップＳＣ－７）、ｉが最終であるかを判定し（ステップＳＣ－８）、最終になるまでステップＳＣ－３からステップＳＣ－７についてループ処理を行う。
再び、図１に戻り、本発明は、各測定対象の遺伝子ＲＮＡについて抽出されたシグナル強度に基づいてｓｃａｔｔｅｒｅｄ　ｐｌｏｔ等を作成し、データマイニングやクラスタリングなどのデータ解析処理を行う（ステップＳＡ－４）。
［システム全体構成］
次に、図５から図７等を参照してＤＮＡマイクロアレイデータ処理システム（以下、「本システム」という。）のシステム構成について説明する。図５は、本発明が適用される本システムの構成の一例を示すブロック図であり、該構成のうち本発明に関係する部分のみを概念的に示している。
図５に示すように、本システムは、概略的に、ＤＮＡマイクロアレイの発光支持体のスキャニングを行う共焦点レーザー顕微鏡型スキャナ装置等のスキャナ装置２００（ここで、スキャナ装置２００と顕微鏡とを組み合わせたシステムを「ＤＮＡマイクロアレイスキャニングシステム」という。）と、当該スキャナ装置２００によりスキャニングされたスキャニング画像について情報処理を行うＤＮＡマイクロアレイデータ処理装置１００とを含

んで構成される。
本システムにおいて、スキャナ装置２００は、ＤＮＡマイクロアレイを標本焦点面に取付した顕微鏡、および、ＤＮＡマイクロアレイデータ処理装置１００に接続され、顕微鏡の画像をスキャニングしてスキャニング画像のデータをＤＮＡマイクロアレイデータ処理装置１００に送信する機能を有する。
また、ＤＮＡマイクロアレイデータ処理装置１００は、スキャナ装置２００から取得したスキャニング画像等のデータについて画像解析等の情報処理を行う機能を有する。
以下にＤＮＡマイクロアレイスキャニングシステムおよびＤＮＡマイクロアレイデータ処理装置１００の構成について順に詳細に説明する。
［システム構成－ＤＮＡマイクロアレイスキャニングシステム］
まず、ＤＮＡマイクロアレイスキャニングシステムの原理構成について説明する。図７は、図５に示すＤＮＡマイクロアレイスキャニングシステムの機能ブロックの一例を示す図であり、ＤＮＡマイクロアレイスキャニングシステムの機能のうち本発明に関係する部分のみを概念的に示している。なお、図７においては、共焦点レーザー顕微鏡型スキャナ装置を用いる場合を一例に説明するが、本発明はこの場合に限定されるものではなく、他の実施の形態においては、他のいずれのスキャナ装置を用いる場合においても適用することができる。
図７においてＤＮＡマイクロアレイスキャニングシステムは、概略的に、複数の励起波長のレーザー光線を照射するレーザー光源部２０２、励起光と蛍光の分離を行うビームスプリッター２０４、放射された蛍光を集光する対物レンズ２０６、光の方向を変えるミラー２０８、蛍光色素の蛍光ピーク付近の狭い範囲の波長を通しそれ以外の波長の励起光などをカットする干渉フィルターとして機能するエミッションフィルター２１０、共焦点性を実現するための受光器レンズ２１２と共焦点ピンホール２１４、および、ＰＭＴやＣＣＤ等により低レベルの光を電気信号に変換する受光器２１６を備えて構成されている。
次に図７を参照してＤＮＡマイクロアレイスキャニングシステムの原理の概要について説明する。レーザー光源部２０２により照射された所定の蛍光色素を蛍光発色させる平行レーザー光線（励起光線）は、ビームスプリッター２０４により反射されて対物レンズ２０６の中に入る。レーザー光線はＤＮＡマイクロアレイの基板上で焦点を結び、その点から蛍光が全ての方向に球状に放射される。
またレーザー光線は反射されて逆行し、受光器２１６の方向に向かう。球状に放射された蛍光の一部が対物レンズ２０６によって集光され、平行ビームに変換される。対物レンズ２０６は反射されたレーザー光も集光する。このレーザー光はビームスプリッター２０４に当たって反射され、そのほとんどは逆行してレーザー光源部２０２に向かう。蛍光のほとんどはビームスプリッター２０４を通過して受光器２１６へ向かう。
ビームスプリッター２０４を通過した光はミラー２０８によってその方向が変えられ、エミッションフィルター２１０へと向かう。エミッションフィルター２１０で狭い蛍光バンドが選択され、残っているレーザー光の全てが遮断される。
共焦点性は受光器レンズ２１２と共焦点ピンホール２１４によって実現される。すなわち、蛍光の平行ビームは受光器レンズ２１２によって収束し、小さな直径となる。共焦点ピンホール２１４の位置で焦点を結んだ光線は共焦点ピンホール２１４を通過し、それ以外の光はブロックされる。また共焦点ピンホール２１４の前後で焦点を結ぶ光は遮断され、対物レンズ２０６の焦点深度を制限する効果をもつ。
受光器２１６は例えば光電子倍増管（ＰＭＴ）を用いて、光子を約百万倍の増幅効率で電子に変換する。そして既知の基板移動式スキャナ等（図示せず）により発光支持体のスキャニングを行い、スキャニング画像をＤＮＡマイクロアレイデータ処理装置１００に対して送信する。
［システム構成－ＤＮＡマイクロアレイデータ処理装置１００］
最初に、本システムのＤＮＡマイクロアレイデータ処理装置１００の構成について説明する。図６は、図５に示すＤＮＡマイクロアレイデータ処理装置１００の機能ブロックの一例を示す図であり、ＤＮＡマイクロアレイデータ処理装置１００の機能のうち本発明に関

係する部分のみを概念的に示している。
図６においてＤＮＡマイクロアレイデータ処理装置１００は、概略的に、ＤＮＡマイクロアレイデータ処理装置１００の全体を統括的に制御するＣＰＵ等の制御部１０２、スキャナ装置２００との通信を行う通信インターフェース機能を有する通信制御インターフェース部１０４、入力装置１１２や出力装置１１４に接続される入出力制御インターフェース部１０８、および、各種のデータベースやテーブルなどを格納する記憶部１０６を備えて構成されており、これら各部は任意の通信路を介して通信可能に接続されている。さらに、このＤＮＡマイクロアレイデータ処理装置１００は、有線または無線の通信回線を介して、スキャナ装置２００に通信可能に接続されている。
図６において記憶部１０６に格納される各種のデータベースやテーブル（スキャニング画像データベース１０６ａ～補正用遺伝子測定結果データベース１０６ｃ）は、固定ディスク装置等のストレージ手段であり、各種処理に用いる各種のプログラムやテーブルやファイルやデータベースやウェブページ用ファイル等を格納する。
これら記憶部１０６の各構成要素のうち、スキャニング画像データベース１０６ａは、スキャニング画像やその関連情報を格納したデータベースである。ここで、関連情報は、ＤＮＡマイクロアレイの試験データ、スキャニング画像番号、スキャニング画像数、使用したスパイクＲＮＡ、スキャニング時にコントロールとして感度調整に用いたスパイクＲＮＡなどに関する情報を含んで構成され、画像データ、テキストファイル、音声ファイル、動画ファイル等が含まれてもよい。
また、処理結果データ１０６ｂは、処理結果に関する情報等を格納する処理結果データ格納手段である。
また、補正用遺伝子測定結果データベース１０６ｃは、様々な補正用遺伝子のＤＮＡマイクロアレイによる測定結果に関するデータベースであって、後述する補正用遺伝子補正部１０２ｉが対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子の測定結果を当該データベースから選択して遺伝子リストを作成し、この補正用遺伝子リストに対応する仮代表値を利用して回帰式を求め全ての仮代表値を補正する際に利用するために、各補正用遺伝子の測定結果を格納する補正用遺伝子測定結果データベースである。
また、図６において、通信制御インターフェース部１０４は、ＤＮＡマイクロアレイデータ処理装置１００とスキャナ装置２００（またはルータ等の通信装置を介して通信を行ってもよい）との間における通信制御を行う。すなわち、通信制御インターフェース部１０４は、他の装置と通信回線を介してデータを通信する機能を有する。
また、図６において、入出力制御インターフェース部１０８は、入力装置１１２や出力装置１１４の制御を行う。ここで、出力装置１１４としては、モニタ（家庭用テレビを含む）の他、スピーカを用いることができる（なお、以下においては出力装置１１４をモニタとして記載する場合がある）。また、入力装置１１２としては、キーボード、マウス、および、マイク等を用いることができる。また、モニタも、マウスと協働してポインティングデバイス機能を実現する。
また、図６において、制御部１０２は、ＯＳ（Ｏｐｅｒａｔｉｎｇ　Ｓｙｓｔｅｍ）等の制御プログラム、各種の処理手順等を規定したプログラム、および所要データを格納するための内部メモリを有し、これらのプログラム等により、種々の処理を実行するための情報処理を行う。制御部１０２は、機能概念的に、スキャニング画像取得部１０２ａ、ノーマライズ部１０２ｂ、シグナル強度選択部１０２ｃ、および、データ解析部１０２ｄを備えて構成されている。
このうち、スキャニング画像取得部１０２ａは、複数の遺伝子から調整したコントロール用ＲＮＡおよび／または複数の人工的に合成したコントロールＲＮＡを測定対象のＲＮＡに混入して調整した試料を用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得手段、および、測定対象サンプルに含まれ、かつ測定対象遺伝子以外から選択した複数の遺伝子のＤＮＡをコントロール用スポットとしてスポットしたＤＮＡマイクロアレイを用いてハイブリダイゼーションを行ったＤＮＡマイクロアレイの発光支持体

のスキャニングを行うスキャナ装置からスキャニング画像を取得するスキャニング画像取得手段である。
また、ノーマライズ部１０２ｂは、スキャニング画像取得手段にて取得されたスキャニング画像の複数の遺伝子から調整したコントロール用ＲＮＡおよび／または人工的に合成したコントロール用ＲＮＡが結合した各コントロール用スポットのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションをコントロール用ＲＮＡ毎に実行するノーマライズ手段、および、スキャニング画像取得手段にて取得されたスキャニング画像の複数のコントロール用スポットに結合したＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度になるように、測定対象のＲＮＡが結合した各スポットのシグナル強度のノーマライゼーションを上記コントロール用スポットの遺伝子毎に実行するノーマライズ手段である。
また、ノーマライズ部１０２ｂは、図１２に示すように、統合化部１０２ｅ、棄却値選定部１０２ｆ、仮代表値決定部１０２ｇ、代表値決定部１０２ｊを含んで構成される。ここで、図１２は、ノーマライズ部１０２ｂの構成の一例を示すブロック図である。
図１２の統合化部１０２ｅは、複数の異なるコントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なるコントロール用ＲＮＡの中から選択した一種類または数種類のコントロール用スポットのシグナル強度に換算することにより統合する統合化手段である。
また、棄却値選定部１０２ｆは、統合化部１０２ｅにて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する棄却値選定手段である。
また、仮代表値決定部１０２ｇは、棄却値選定部１０２ｆにより選定された棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の仮平均値を得て各遺伝子の仮代表値を決定する仮代表値決定手段である。
また、代表値決定部１０２ｊは、仮代表値より回帰式を求めて補正して代表値を決定する手段である。ここで、図１３は、代表値決定部１０２ｊの構成の一例を示すブロック図である。図１３に示すように、代表値決定部１０２ｊは、回帰式補正部１０２ｈおよび補正用遺伝子補正部１０２ｉを含んで構成される。
図１３の回帰式補正部１０２ｈは、仮代表値決定部１０２ｇにて決定された仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正する回帰式補正手段である。
また、補正用遺伝子補正部１０２ｉは、あらかじめ補正用遺伝子測定結果データベース１０６ｃに格納された各補正用遺伝子の測定結果に基づいて、対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子のリストを作成し、当該補正用遺伝子リストに対応する仮代表値を選択し、対応する仮代表値を利用して回帰式を求め、全ての仮代表値を補正することにより、対照検体と測定検体の測定上の誤差を補正して代表値を決定する補正用遺伝子補正手段である。
再び図６に戻り、シグナル強度選択部１０２ｃは、ノーマライズ手段にて各コントロール用ＲＮＡによってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段、および、ノーマライズ手段にて各コントロール用スポットの遺伝子によってノーマライズされた各測定対象スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択するシグナル強度選択手段である。
また、データ解析部１０２ｄは、各測定対象の遺伝子ＲＮＡについて抽出されたシグナル強度に基づいてｓｃａｔｔｅｒｅｄ　ｐｌｏｔ等を作成し、データマイニングや階層クラスタリング、Ｋ－平均法、サーポートベクターマシーン（ＳＶＭ）、自己組織化マップ（ＳＯＭ）、マハラノビス距離を用いたクラスタリングなどの判定アルゴリズムデータ解析処理を行うデータ解析手段である。

なお、これら各部によって行なわれる処理の詳細については、後述する。
［システムの処理］
次に、このように構成された本実施の形態における本システムの処理の一例について、以下に図８～図１１等を参照して詳細に説明する。
［スキャニング処理］
まず、スキャナ装置２００等におけるスキャニング処理の詳細について図８等を参照して説明する。図８は、本実施形態における本システムのスキャニング処理の一例を示すフローチャートである。
スキャナ装置２００は、スパイク毎に感度調整を行なわない場合においては（ステップＳＤ－１）、ハイブリダイゼーションを行ったＤＮＡマイクロアレイについてスキャナ装置２００を用いて発光支持体をスキャニングする（ステップＳＤ－２）。
一方、スパイク毎に感度調整を行う場合においては（ステップＳＤ－１）、複数のスパイクＲＮＡのシグナル強度が予め定めた測定シグナル強度（例えば上限値の２分の１等）になるように感度を調整してスパイクＲＮＡ毎に発光支持体をスキャニングする（ステップＳＤ－３）。
ここで、感度調整は、例えば、励起光の強さを変えることによって行ってもよい。すなわち、レーザー光源部２０２において照射される励起光（レーザー光線）の出力量を調整してもよく、また、スキャナ装置２００に可変レーザー減衰器（図示せず）をさらに備えて構成することにより、レーザー光源部２０２から照射される励起光の強さを少なくとも１００：１のレンジで調整してもよい。
また、感度調節は、例えば、受光器２１６による受光感度を調節することにより行ってもよい。すなわち、例えば受光器２１６の光電子倍増管（ＰＭＴ）にかかる電圧を変えてＰＭＴの感度を調整する。
そして、スキャナ装置２００は、スキャニングしたスキャニング画像のデータをＤＮＡマイクロアレイデータ処理装置１００に対して送信する（ステップＳＤ－４）。
これにて、スキャニング処理が終了する。
［スキャニング画像取得処理］
次に、ＤＮＡマイクロアレイデータ処理装置１００におけるスキャニング画像取得処理の詳細について図９等を参照して説明する。図９は、本実施形態における本システムのスキャニング画像取得処理の一例を示すフローチャートである。
まず、ＤＮＡマイクロアレイデータ処理装置１００は、一つのＤＮＡマイクロアレイについて複数のスキャニング画像が存在する場合（すなわち、ステップＳＤ－３において上述したようにスパイク毎に感度調節を行ってスキャニング画像を撮像した場合）には（ステップＳＥ－１）、スキャニング画像取得部１０２ａの処理により、通信制御インターフェース部１０４を介してスキャナ装置２００から取得した複数のスキャニング画像をスキャニング画像データベース１０６ａに格納する（ステップＳＥ－２およびステップＳＥ－４）。
一方、ＤＮＡマイクロアレイデータ処理装置１００は、一つのＤＮＡマイクロアレイについて複数のスキャニング画像が存在しない場合（すなわち、スパイク毎に感度調節を行ってスキャニング画像を撮像しなかった場合）には（ステップＳＥ－１）、スキャニング画像取得部１０２ａの処理により、通信制御インターフェース部１０４を介してスキャナ装置２００から取得した一つのスキャニング画像をスキャニング画像データベース１０６ａに格納する（ステップＳＥ－３およびステップＳＥ－４）。
ここで、スキャニング画像取得部１０２ａは、各ＤＮＡマイクロアレイと対応付けてＤＮＡマイクロアレイ番号、ＤＮＡマイクロアレイの実験データ、撮像したスキャニング画像数、スキャニング画像番号、および、使用したスパイクＲＮＡ等に関する情報をスキャニング画像データベース１０６ａに格納し、さらに、各スキャニング画像データと対応付けてスキャニング画像番号、および、スキャニング時にコントロールとして感度調整に用いたスパイクＲＮＡ等に関する情報を、スキャニング画像データベース１０６ａに格納する。

これにて、スキャニング画像取得処理が終了する。
［ノーマライズ処理］
次に、ＤＮＡマイクロアレイデータ処理装置１００におけるノーマライズ処理の詳細について図１０等を参照して説明する。図１０は、本実施形態における本システムのノーマライズ処理の一例を示すフローチャートである。
まず、ＤＮＡマイクロアレイデータ処理装置１００は、ノーマライズ部１０２ｂの処理により、スキャニング画像データベース１０６ａにアクセスして、所望のＤＮＡマイクロアレイについて撮像したスキャニング画像数を読み込む（ステップＳＦ－１）。
そして、ノーマライズ部１０２ｂは、スキャニング番号を示す添え字ｎを１に初期化した後（ステップＳＦ－２）、当該ＤＮＡマイクロアレイのｎ番目（最初は１番目）のスキャニング画像をスキャニング画像データベース１０６ａから読み込む（ステップＳＦ－３）。
そして、ノーマライズ部１０２ｂは、図３のステップＳＢ－２～ステップＳＢ－６で上述した処理を実行することにより、ｎ番目のスキャニング画像についてノーマライゼーションを行う（ステップＳＦ－４）。すなわち、ノーマライズ部１０２ｂは、スキャニング画像の複数のスパイクＲＮＡのシグナル強度がそれぞれ予め定めた基準シグナル強度（例えば、上限値の２分の１等）になるように、測定対象のＲＮＡの各スポットのシグナル強度のノーマライゼーションをスパイクＲＮＡ毎に実行する。
そして、ノーマライズ部１０２ｂは、ｎに１を加算した後（ステップＳＦ－５）、ステップＳＦ－１にて取得したスキャニング画像数とｎとを比較することによりｎが最終であるかを判定し（ステップＳＦ－６）、ｎが最終になるまでステップＳＦ－３からステップＳＦ－６についてループ処理を行う。
そして、ノーマライズ部１０２ｂは、全てのスキャニング画像についてノーマライズ後の測定データを処理結果データ１０６ｂに格納する（ステップＳＦ－７）。
これにて、ノーマライズ処理が終了する。
［シグナル強度選択処理］
次に、ＤＮＡマイクロアレイデータ処理装置１００におけるシグナル強度選択処理の詳細について図１１等を参照して説明する。図１１は、本実施形態における本システムのシグナル強度選択処理の一例を示すフローチャートである。
まず、ＤＮＡマイクロアレイデータ処理装置１００は、シグナル強度選択部１０２ｃの処理により、処理結果データ１０６ｂにアクセスしてＤＮＡマイクロアレイに関するノーマライズ後の測定データを全て読み込む（ステップＳＧ－１）。
そして、シグナル強度選択部１０２ｃは、図４を用いて上述したように、各スパイクＲＮＡによってノーマライズされた各スポットのシグナル強度に基づいて、スキャニング画像の各測定対象のＲＮＡのシグナル強度のうち最適なシグナル強度を選択する（ステップＳＧ－２）。
そして、ＤＮＡマイクロアレイデータ処理装置１００は、データ解析部１０２ｄの処理により、選択されたシグナル強度に基づいて、各種のプログラムを用いて所望のデータ解析を行う。
これにて、シグナル強度選択処理が終了する。
［本発明のノーマライズ処理等の別の実施形態］
次に、ＤＮＡマイクロアレイデータ処理装置１００におけるノーマライズ処理等の別の実施形態の詳細について図１４および図１５等を参照して説明する。図１４は、本実施形態における本システムのノーマライズ処理などの別の実施形態の一例を示すフローチャートである。
まず、異なったスキャン条件で取得したデータの統合化のための前処理として、上述したスキャナ装置２００によるスキャニング条件の設定を行う（ステップＳＨ－１）。
ここで、図１４のＳＨ－１の部分の詳細を示した図１５は、スキャナ装置２００によるスキャニング条件の設定の一例を示すフローチャートである。図１５に示すように、配置したコントロール用スポットの取込値（シグナル強度）が一番安定したシグナルとして取り

込めるような条件で、コントロール毎に感度調整を行い、スキャン条件を１つまたは複数決定する（ステップＳＪ－１）。
ここで、コントロール用スポットの取込値（シグナル強度）は、スキャナ装置２００の上限値の半分程度に設定することが好ましい。
そして、ステップＳＪ－１により決定されたスキャン条件により、発現量に応じてスキャナ装置２００によりスキャニングを行い、各スキャニング画像データを取得する（ステップＳＪ－２～ステップＳＪ－６）。
そして、スキャン条件に基づいて、シグナル強度の低限界値（下限）と高限界値（上限）を選択する（ステップＳＪ－７）。
再び図１４に戻り、ＤＮＡマイクロアレイデータ処理装置１００（図６）は、スキャニング画像取得処理１０２ａ（図６）によって、ステップＳＨ－１においてスキャナ装置２００が取得した各スキャニング画像データを取得すると、統合化部１０２ｅ（図１２）の処理により、複数の異なるコントロール用ＲＮＡを用いたスキャニングにより複数の異なる条件で取得した各スポットのシグナル強度を、当該複数の異なるコントロール用ＲＮＡの中から選択した一種類または数種類のコントロール用スポットのシグナル強度に換算することにより統合する（ステップＳＨ－２）。
ここで、統合化部１０２ｅ（図１２）による異なった条件でスキャンニングしたデータの統合化は、複数のスパイクＲＮＡの中から一種類または数種類の選択した基準遺伝子のスキャンニングデータを用いて数値の最適化を行うことによりシグナル強度を換算する。統合化部１０２ｅ（図１２）は、同一スポットの異なったスキャンニング条件での最適化数値より、統合の代表値を求める。代表値の決定法は、例えば平均値、中央値等とする。
そして、ＤＮＡマイクロアレイデータ処理装置１００（図６）は、棄却値選定部１０２ｆ（図１２）の処理により、ステップＳＨ－２にて換算した各スポットのシグナル強度または二色競合法における対照検体と測定検体のシグナル強度の比を同一遺伝子のスポット間で比較し、棄却検定またはユークリッド距離に基づき棄却値を選定する（ステップＳＨ－３）。
ここで、棄却値選定部１０２ｆ（図１２）による棄却値の選択は、例えば次の１）～３）を単独または複合で使用することができる。
１）同一遺伝子のＣｙ３値またはＣｙ５値での棄却検定を行う。
２）同一遺伝子のＣｙ３値／Ｃｙ５値での棄却検定を行う。
３）同一遺伝子間のＣｙ３値およびＣｙ５値のユークリッド距離を求めて棄却値を選定する。
なお、棄却値選定部１０２ｆ（図１２）による棄却計算方法はマイクロアレイチップ性能、目的により選択可能とする。
そして、ＤＮＡマイクロアレイデータ処理装置１００（図６）は、仮代表値決定部１０２ｇ（図１２）の処理により、ステップＳＨ－３により選定された棄却値を用いて棄却した後の残された同一遺伝子スポットのシグナル強度の平均値または中央値を得て各遺伝子の仮代表値を決定する（ステップＳＨ－４）。
ここで、仮代表値決定部１０２ｇ（図１２）による、選択した棄却値を除いた仮代表値の算出は、マイクロアレイの実験条件・精度・スポット数などにより、例えば次の１）～４）を自由に選択できる。
１）同一遺伝子の棄却値を除いてＣｙ３値またはＣｙ５値の個々の平均値を求めて仮代表値とする。
２）同一遺伝子の棄却値を除いてＣｙ３値またはＣｙ５値の個々の中央値を求めて仮代表値とする。
３）同一遺伝子のＣｙ３値／Ｃｙ５値（比）の平均値を求めて仮代表値とする。
４）同一遺伝子のＣｙ３値／Ｃｙ５値（比）の中央値を求めて仮代表値とする。
そして、ＤＮＡマイクロアレイデータ処理装置１００（図６）は、仮代表値決定部１０２ｇ（図１２）の処理により、Ｃｙ３値／Ｃｙ５値の比の値を代表値としない場合には（ステップＳＨ－５）、仮代表値のＣｙ３値およびＣｙ５値仮代表値の算出を代表値決定部１

０２ｊ（図１２）の処理により代表値を決定する（ステップＳＨ－６、代表値決定部１０２ｊ（図１２））。
ここで、仮代表値決定部１０２ｇ（図１２）および代表値決定部１０２ｊ（図１２）によるステップＳＨ－５およびＳＨ－６における代表値の決定は、マイクロアレイの実験条件・精度・スポット数などにより、例えば次の１）または２）を自由に選択できる。
１）統合・棄却後のＣｙ３値およびＣｙ５値の仮代表値を代表値とする。（ＳＨ－５から直接ＳＨ－７に移行する）
２）統合・棄却後のＣｙ３値およびＣｙ５値を内部の補正用遺伝子を使用して補正して代表値とする。（ＳＨ－５からＳＨ－６（代表値決定部１０２ｊ（図１２））を経由してＳＨ－７に移行する）
ここで、補正手法として、代表値決定部１０２ｊ（図１３）は、回帰式補正部１０２ｈ（図１３）の処理により、仮代表値決定部１０２ｇにて決定した仮代表値から求めた回帰式により、各スポットのシグナル強度が直線性を持つように補正してもよい。
また、代表値決定部１０２ｊ（図１３）は、補正用遺伝子補正部１０２ｉ（図１３）の処理により、あらかじめ補正用遺伝子測定結果データベース１０６ｃに登録された各補正用遺伝子の測定結果に従って対照検体と測定検体で同等のシグナル強度を示す補正用遺伝子のリストを選択し、当該補正用遺伝子リストに対応する仮代表値を選択し、この補正用遺伝子リストに対応する仮代表値を利用して回帰式を求めて全ての仮代表値を補正することにより、対照検体と測定検体の測定上の誤差を補正して代表値を決定して測定上の誤差を補正してもよい。
そして、ＤＮＡマイクロアレイデータ処理装置１００（図６）は、データ解析部１０２ｄ（図６）の処理により、Ｋ－平均法や階層クラスタリング手法（例えば、Ｓｔｅｅｎ Ｋｎｕｄｓｅｎ著「ＤＮＡマイクロアレイデータ解析入門」（羊土社、２００２年）などを参照。）、サポートベクターマシーン（ＳＶＭ）を用いたクラスタリング手法（例えば、Ｉｉｚｕｋａ Ｎら”Ｏｌｉｇｏｎｕｃｌｅｏｔｉｄｅ ｍｉｃｒｏａｒｒａｙ ｆｏｒ ｐｒｅｄｉｃｔｉｏｎ ｏｆ ｅａｒｌｙ ｉｎｔｒａｈｅｐａｔｉｃ ｒｅｃｕｒｒｅｎｃｅ ｏｆ ｈｅｐａｔｏｃｅｌｌｕｌａｒ ｃａｒｃｉｎｏｍａ ａｆｔｅｒ ｃｕｒａｔｉｖｅ ｒｅｓｅｃｔｉｏｎ．”（Ｌａｎｃｅｔ ２００３ Ｍａｒ １５；３６１（９３６１）：９２３－９）、Ｖａｌｅｎｔｉｎｉ Ｇ．”Ｇｅｎｅ ｅｘｐｒｅｓｓｉｏｎ ｄａｔａ ａｎａｌｙｓｉｓ ｏｆ ｈｕｍａｎ ｌｙｍｐｈｏｍａ ｕｓｉｎｇ ｓｕｐｐｏｒｔ ｖｅｃｔｏｒ ｍａｃｈｉｎｅｓ ａｎｄ ｏｕｔｐｕｔ ｃｏｄｉｎｇ ｅｎｓｅｍｂｌｅｓ．”（Ａｒｔｉｆ Ｉｎｔｅｌｌ Ｍｅｄ ２００２ Ｎｏｖ；２６（３）：２８１－３０４）、Ｐａｖｌｉｄｉｓ Ｐら”Ｌｅａｒｎｉｎｇ ｇｅｎｅ ｆｕｎｃｔｉｏｎａｌ ｃｌａｓｓｉｆｉｃａｔｉｏｎｓ ｆｒｏｍ ｍｕｌｔｉｐｌｅ ｄａｔａ ｔｙｐｅｓ．”（ＪＣｏｍｐｕｔ Ｂｉｏｌ ２００２；９（２）：４０１－１１）、Ｆｕｒｅｙ ＴＳら”Ｓｕｐｐｏｒｔ ｖｅｃｔｏｒ ｍａｃｈｉｎｅ ｃｌａｓｓｉｆｉｃａｔｉｏｎ ａｎｄ ｖａｌｉｄａｔｉｏｎ ｏｆ ｃａｎｃｅｒ ｔｉｓｓｕｅ ｓａｍｐｌｅｓｕｓｉｎｇ ｍｉｃｒｏａｒｒａｙ ｅｘｐｒｅｓｓｉｏｎ ｄａｔａ．”（Ｂｉｏｉｎｆｏｒｍａｔｉｃｓ ２０００ Ｏｃｔ；１６（１０）：９０６－１４）、などを参照。）、自己組織化マップ（ＳＯＭ）を用いたクラスタリング手法（例えば、Ｌａｍｅｎｄｏｌａ ＤＥら”Ｍｏｌｅｃｕｌａｒ Ｄｅｓｃｒｉｐｔｉｏｎ ｏｆ Ｅｖｏｌｖｉｎｇ Ｐａｃｌｉｔａｘｅｌ Ｒｅｓｉｓｔａｎｃｅ ｉｎ ｔｈｅ ＳＫＯＶ－３ Ｈｕｍａｎ Ｏｖａｒｉａｎ Ｃａｒｃｉｎｏｍａ Ｃｅｌｌ Ｌｉｎｅ．”（Ｃａｎｃｅｒ Ｒｅｓ ２００３ Ｍａｙ １；６３（９）：２２００－５）、Ｈｅｒｒｅｒｏ Ｊら”Ｃｏｍｂｉｎｉｎｇ ｈｉｅｒａｒｃｈｉｃａｌ ｃｌｕｓｔｅｒｉｎｇ ａｎｄ ｓｅｌｆ－ｏｒｇａｎｉｚｉｎｇ ｍａｐｓ ｆｏｒ ｅｘｐｌｏｒａｔｏｒｙａｎａｌｙｓｉｓ ｏｆ ｇｅｎｅ ｅｘｐｒｅｓｓｉｏｎ ｐａｔｔｅｒｎｓ．”（Ｊ Ｐｒｏｔｅｏｍｅ Ｒｅｓ ２００２ Ｓｅｐ－Ｏｃｔ；１（５）：４６７－７０）、Ｘｉａｏ Ｌら”Ｃｏｍｐｏｎｅｎｔ ｐｌａｎｅ ｐｒｅｓｅｎｔａｔｉｏｎ ｉｎｔｅｇｒａｔｅｄ ｓｅｌｆ－ｏｒｇａｎｉｚｉｎｇ ｍａｐ ｆｏ

r microarray dataanalysis.”（FEBS Lett 2003 Mar 13；538（1－3）：117－24）、Wang Jら”Clustering of the SOM easily reveals distinct gene expression patterns：resultsof a reanalysis of lymphoma study.”（BMC Bioinformatics 2002 Nov 24；3（1）：36）、Toronen Pら”Analysis of gene expression data using self－organizing maps.”（FEBS Lett 1999 May 21；451（2）：142－6）、などを参照。）、マハラノビス距離を用いたクラスタリング手法（例えば、Chilingaryan Aら”Multivariate approach for selecting sets of differentially expressed genes.”（Math Biosci 2002 Mar；176（1）：59－69）、などを参照。）等による各種の判定アルゴリズムにより解析を行う。（ステップSH－7、データ解析部102d（図6））。

そして、DNAマイクロアレイデータ処理装置100（図6）は、ノーマライズ部102b（図6）の処理結果およびデータ解析部102d（図6）により、ノーマライズデータと最終結果を取得し、記憶部106（図6）の処理結果データ106b（図6）の所定の記憶領域に格納する。

［他の実施の形態］

さて、これまで本発明の実施の形態について説明したが、本発明は、上述した実施の形態以外にも、上記特許請求の範囲に記載した技術的思想の範囲内において種々の異なる実施の形態にて実施されてよいものである。

例えば、DNAマイクロアレイデータ処理装置100がスタンドアローンの形態で処理を行う場合を一例に説明したが、DNAマイクロアレイデータ処理装置100とは別筐体で構成されるクライアント端末からの要求に応じて処理を行い、その処理結果を当該クライアント端末に返却するように構成してもよい。

また、上記の実施形態においては、コントロール用遺伝子としてスパイク用遺伝子を用いてノーマライズを行う場合を一例に説明したが、本発明はかかる場合に限定されるものではなく、試料内に含まれる遺伝子で、細胞毎の発現の変動が少ない遺伝子（例えば、beta－actin、GAPDH、poly（A） binding protein（PABP）、glycyl tRNA synthetase、hnRNP H1等）等を複数用いた場合にも同様に適用することができる。

また、DNAマイクロアレイデータ処理装置100が、スキャナ装置200やDNAマイクロアレイスキャニングシステムの動作を制御する機能を備えていてもよい。すなわち、DNAマイクロアレイデータ処理装置100が送信する命令に従って、スキャナ装置200における感度調整などの動作を実行してもよい。また、スキャナ装置200やDNAマイクロアレイスキャニングシステムの動作を制御する機能を備えた別のコンピュータ装置を用いてもよい。

また、実施形態において説明した各処理のうち、自動的に行なわれるものとして説明した処理の全部または一部を手動的に行うこともでき、あるいは、手動的に行なわれるものとして説明した処理の全部または一部を公知の方法で自動的に行うこともできる。

この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種の登録データや検索条件等のパラメータを含む情報、画面例、データベース構成については、特記する場合を除いて任意に変更することができる。

また、DNAマイクロアレイデータ処理装置100に関して、図示の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。

例えば、DNAマイクロアレイデータ処理装置100の各部または各装置が備える処理機能、特に制御部102にて行なわれる各処理機能については、その全部または任意の一部を、CPU（Central Processing Unit）および当該CPUにて解釈実行されるプログラムにて実現することができ、あるいは、ワイヤードロジックによ

るハードウェアとして実現することも可能である。なお、プログラムは、後述する記録媒体に記録されており、必要に応じてＤＮＡマイクロアレイデータ処理装置１００に機械的に読み取られる。
すなわち、ＲＯＭまたはＨＤなどの記憶部１０６などには、ＯＳ（Ｏｐｅｒａｔｉｎｇ　Ｓｙｓｔｅｍ）と協働してＣＰＵに命令を与え、各種処理を行うためのコンピュータプログラムが記録されている。このコンピュータプログラムは、ＲＡＭ等にロードされることによって実行され、ＣＰＵと協働して制御部１０２を構成する。また、このコンピュータプログラムは、ＤＮＡマイクロアレイデータ処理装置１００に対して任意のネットワークを介して接続されたアプリケーションプログラムサーバに記録されてもよく、必要に応じてその全部または一部をダウンロードすることも可能である。
また、本発明にかかるプログラムを、コンピュータ読み取り可能な記録媒体に格納することもできる。ここで、この「記録媒体」とは、フレキシブルディスク、光磁気ディスク、ＲＯＭ、ＥＰＲＯＭ、ＥＥＰＲＯＭ、ＣＤ－ＲＯＭ、ＭＯ、ＤＶＤ等の任意の「可搬用の物理媒体」や、各種コンピュータシステムに内蔵されるＲＯＭ、ＲＡＭ、ＨＤ等の任意の「固定用の物理媒体」、あるいは、ＬＡＮ、ＷＡＮ、インターネットに代表されるネットワークを介してプログラムを送信する場合の通信回線や搬送波のように、短期にプログラムを保持する「通信媒体」を含むものとする。
また、「プログラム」とは、任意の言語や記述方法にて記述されたデータ処理方法であり、ソースコードやバイナリコード等の形式を問わない。なお、「プログラム」は必ずしも単一的に構成されるものに限られず、複数のモジュールやライブラリとして分散構成されるものや、ＯＳ（Ｏｐｅｒａｔｉｎｇ　Ｓｙｓｔｅｍ）に代表される別個のプログラムと協働してその機能を達成するものをも含む。なお、実施の形態に示した各装置において記録媒体を読み取るための具体的な構成、読み取り手順、あるいは、読み取り後のインストール手順等については、周知の構成や手順を用いることができる。
記憶部１０６に格納される各種のデータベース等（スキャニング画像データベース１０６ａ～補正用遺伝子測定結果データベース１０６ｃ）は、ＲＡＭ、ＲＯＭ等のメモリ装置、ハードディスク等の固定ディスク装置、フレキシブルディスク、光ディスク等のストレージ手段であり、各種処理やウェブサイト提供に用いる各種のプログラムやテーブルやファイルやデータベースやウェブページ用ファイル等を格納する。
また、ＤＮＡマイクロアレイデータ処理装置１００は、既知のパーソナルコンピュータ、ワークステーション等の情報処理端末等の情報処理装置にプリンタやモニタやイメージスキャナ等の周辺装置を接続し、該情報処理装置に本発明の方法を実現させるソフトウェア（プログラム、データ等を含む）を実装することにより実現してもよい。
さらに、ＤＮＡマイクロアレイデータ処理装置１００の分散・統合の具体的形態は図示のものに限られず、その全部または一部を、各種の負荷等に応じた任意の単位で、機能的または物理的に分散・統合して構成することができる。例えば、各データベースを独立したデータベース装置として独立に構成してもよく、また、処理の一部をＣＧＩ（Ｃｏｍｍｏｎ　Ｇａｔｅｗａｙ　Ｉｎｔｅｒｆａｃｅ）を用いて実現してもよい。
［実施例］
以下に本発明を実施例に基づき具体的に説明する。もっとも本発明は下記実施例に限定されるものではない。
１．複数の人工（コントロール）ＲＮＡスポットのシグナル強度を基準としたデータの取込み
コントロール用ＲＮＡをウミシイタケルシフェラーゼ遺伝子（Ｒｌｕｃ、表中ではＲＬ）、バキュロウイルス膜蛋白遺伝子（ｇｐ６４、表中ではＧＰ）、及びラムダファージｅａ２２遺伝子（ｅａ２２、表中ではＬＰ）を標識する試料ＲＮＡ（全ＲＮＡ、ｍＲＮＡ、増幅ＲＮＡ）に添加した（上述した図２等参照）。添加量はＣｙ３用、Ｃｙ５用のそれぞれの試料ＲＮＡ６μｇに５０００ｐｇ（Ｒｌｕｃ）、１６００ｐｇ（ｇｐ６４）、３１２．５ｐｇ（ｅａ２２）とした。
常法により検体ＲＮＡをＣｙ３で、対照ＲＮＡをＣｙ５で標識し、ＤＮＡマイクロアレイ

（自家製）上でハイブリダイゼーションを行い、スキャニングの試料とした。
スキャナは、ＳｃａｎＡｒｒａｙ５０００（製品名）（パーキンエルマーライフサイエンスジャパン株式会社（会社名））を使用した。スキャン条件は、まず、人工ＲＮＡとして高濃度で添加したウミホタルルシフェラーゼ遺伝子（以下ＲＬという）のスポットがスキャナ上限値の５０％シグナル値（上記スキャナでは絶対シグナル値として約３００００になる）になる条件でスキャニングした（以下この条件を「ＲＬ５０％」、この条件で得た値を「ＲＬ５０％値」という。ＧＰ、ＬＰでも同じ。）。次に、人工ＲＮＡとして中濃度で添加したバキュロウイスル膜蛋白遺伝子（以下ＧＰという）のスポットがスキャナ上限値の５０％シグナル値（ＧＰ５０％値）になる条件でスキャニングした。最後に、人工ＲＮＡを低濃度で添加したラムダファージｅａ２２遺伝子（以下ＬＰという）のスポットがスキャナ上限値の５０％シグナル値になるようにスキャンニングした。すなわちＲＬを基準としたスキャニングでは弱いレーザーパワーで、ＬＰを基準としたスキャニングでは強いレーザーパワーでデータを取込んだことになる。このようにしてスキャニングしたデータでは、各スポットに対して６つの異なったスキャン条件での結果が得られることになる（Ｃｙ３とＣｙ５についてそれぞれ高・中・低）。
このようにデータを取得することにより、「従来の技術」で述べた測定上の誤差、特に低発現領域と高発現領域で見られる非線形の問題を解消することができる。即ち、この部分の解決策として、レンジ決めのスパイク（コントロール）を使って線形性の高いデータを３種類集め、後に補正によって統合するため、線形性を保った状態のデータを取扱えるというメリットがある。また、取扱うファイルがすべて線形なので、スケール化（ノーマライズ）する前に複数のファイルを１つのファイルに統合して、そのまま取り扱える点にメリットが非常に大きい。
表１にその結果をアレイ上の代表的な遺伝子及びコントロール遺伝子（ＲＬ、ＧＰ、ＬＰ）について示す。なおＣｙ３とＣｙ５のサンプルは同じもので、表中のＣｙ３＿ＧＰ、Ｃｙ５＿ＧＰはＧＰスポットが、Ｃｙ３＿ＲＬ、Ｃｙ５＿ＲＬはＲＬスポットが、Ｃｙ３＿ＬＰ、Ｃｙ５＿ＬＰはＬＰスポットが、それぞれＣｙ３、Ｃｙ５の波長でシグナル強度３００００になるように取り込んだものである。各遺伝子はそれぞれ４スポット、コントロールは７スポットある。表中の値で「６５５３５」は飽和状態を示す。

表１

| | Cy3_RL | Cy5_RL | Cy3_GP | Cy5_GP | Cy3_LP | Cy5_LP |
|---|---|---|---|---|---|---|
| 遺伝子名 | シグナル中央値 | シグナル中央値 | シグナル中央値 | シグナル中央値 | シグナル中央値 | シグナル中央値 |
| alpha-1 antitrypsin | 591 | 279 | 6497 | 5598 | 9114 | 7668 |
| alpha-1 antitrypsin | 343 | 113 | 5335 | 4429 | 7834 | 6208 |
| alpha-1 antitrypsin | 425 | 154 | 5861 | 4235 | 7591 | 5885 |
| alpha-1 antitrypsin | 486 | 341 | 6911 | 5972 | 9670 | 8118 |
| NADH dehydrogenase 4 | 1646 | 1039 | 15744 | 16501 | 21923 | 22495 |
| NADH dehydrogenase 4 | 1792 | 1324 | 18422 | 19623 | 25745 | 26710 |
| NADH dehydrogenase 4 | 1293 | 716 | 14851 | 11803 | 20495 | 15488 |
| NADH dehydrogenase 4 | 1338 | 831 | 15101 | 14357 | 21584 | 19210 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 16S rRNA, mitochondrial | 42279 | 38323 | 65535 | 65535 | 65535 | 65535 |
| 16S rRNA, mitochondrial | 43473 | 41610 | 65535 | 65535 | 65535 | 65535 |
| 16S rRNA, mitochondrial | 41526 | 38726 | 65535 | 65535 | 65535 | 65535 |
| 16S rRNA, mitochondrial | 41948 | 39046 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L41 | 33543 | 25184 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L41 | 40455 | 28095 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L41 | 31935 | 23652 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L41 | 37689 | 26766 | 65535 | 65535 | 65535 | 65535 |
| NADH dehydrogenase | 1743 | 1441 | 18507 | 22304 | 25444 | 29814 |
| NADH dehydrogenase | 2477 | 2007 | 23122 | 28221 | 33374 | 38920 |
| NADH dehydrogenase | 1357 | 1041 | 14375 | 17126 | 19305 | 22473 |
| NADH dehydrogenase | 1843 | 1545 | 19081 | 23413 | 26274 | 31518 |
| Apolipoprotein E | 76 | 0 | 2813 | 1346 | 3441 | 1670 |
| Apolipoprotein E | 31 | 0 | 2670 | 1284 | 3347 | 1755 |
| Apolipoprotein E | 121 | 0 | 3053 | 1322 | 3660 | 1960 |
| Apolipoprotein E | 120 | 0 | 2406 | 1215 | 4235 | 1616 |
| ferritin light chain | 37379 | 29494 | 65535 | 65535 | 65535 | 65535 |
| ferritin light chain | 36148 | 27493 | 65535 | 65535 | 65535 | 65535 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ferritin light chain | 34302 | 26104 | 65535 | 65535 | 65535 | 65535 |
| ferritin light chain | 35709 | 24118 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S12 | 30559 | 28048 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S12 | 29224 | 20825 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S12 | 31400 | 23907 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S12 | 31401 | 25518 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L35 | 10654 | 7871 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L35 | 11584 | 7865 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L35 | 10124 | 7352 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein L35 | 10938 | 6292 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S6 | 36319 | 33782 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S6 | 38607 | 37562 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S6 | 30235 | 27750 | 65535 | 65535 | 65535 | 65535 |
| ribosomal protein S6 | 33729 | 30549 | 65535 | 65535 | 65535 | 65535 |
| kininogen | 137 | 0 | 3465 | 1596 | 4850 | 2394 |
| kininogen | 138 | 0 | 2932 | 1622 | 3739 | 2050 |
| kininogen | 101 | 0 | 3871 | 1805 | 5263 | 2154 |
| kininogen | 109 | 0 | 3333 | 1627 | 4810 | 1950 |

| | | | | | | |
|---|---|---|---|---|---|---|
| GAPDH | 26895 | 24122 | 65535 | 65535 | 65535 | 65535 |
| GAPDH | 24558 | 21632 | 65535 | 65535 | 65535 | 65535 |
| GAPDH | 23979 | 18509 | 65535 | 65535 | 65535 | 65535 |
| GAPDH | 24847 | 15795 | 65535 | 65535 | 65535 | 65535 |
| thymosin beta-10 | 20790 | 17925 | 65535 | 65535 | 65535 | 65535 |
| thymosin beta-10 | 23000 | 18772 | 65535 | 65535 | 65535 | 65535 |
| thymosin beta-10 | 22920 | 17017 | 65535 | 65535 | 65535 | 65535 |
| thymosin beta-10 | 22912 | 16533 | 65535 | 65535 | 65535 | 65535 |
| G6PD | 233 | 0 | 4337 | 2276 | 5313 | 3396 |
| G6PD | 277 | 0 | 4889 | 2713 | 6218 | 3710 |
| G6PD | 211 | 0 | 3999 | 2098 | 5390 | 2808 |
| G6PD | 248 | 0 | 3540 | 1995 | 5790 | 2754 |
| RL (人工RNA) | 33435 | 31642 | 65535 | 65535 | 65535 | 65535 |
| RL (人工RNA) | 31053 | 29761 | 65535 | 65535 | 65535 | 65535 |
| RL (人工RNA) | 33476 | 32309 | 65535 | 65535 | 65535 | 65535 |
| RL (人工RNA) | 32489 | 32457 | 65535 | 65535 | 65535 | 65535 |
| RL (人工RNA) | 30731 | 21094 | 65535 | 65535 | 65535 | 65535 |
| RL (人工RNA) | 35997 | 26235 | 65535 | 65535 | 65535 | 65535 |
| RL (人工RNA) | 42847 | 32467 | 65535 | 65535 | 65535 | 65535 |
| GP (人工RNA) | 3743 | 2629 | 36279 | 37969 | 47742 | 50220 |
| GP (人工RNA) | 2902 | 1951 | 29396 | 29223 | 39148 | 38537 |
| GP (人工RNA) | 2683 | 1948 | 26241 | 27466 | 36118 | 37177 |
| GP (人工RNA) | 2911 | 2142 | 26069 | 30238 | 38064 | 39632 |
| GP (人工RNA) | 3077 | 2349 | 28986 | 30726 | 39314 | 40521 |
| GP (人工RNA) | 3262 | 1981 | 32051 | 28663 | 41841 | 37472 |
| GP (人工RNA) | 4559 | 2937 | 41807 | 39764 | 56294 | 53478 |

| | | | | | | |
|---|---|---|---|---|---|---|
| LP(人工 RNA) | 2488 | 1692 | 25269 | 28592 | 33954 | 37530 |
| LP(人工 RNA) | 2212 | 1676 | 23076 | 25104 | 30601 | 32904 |
| LP(人工 RNA) | 1914 | 1069 | 20014 | 17802 | 28835 | 23300 |
| LP(人工 RNA) | 2030 | 1535 | 20986 | 23041 | 30322 | 30375 |
| LP(人工 RNA) | 2465 | 1783 | 23269 | 25574 | 32294 | 34271 |
| LP(人工 RNA) | 2583 | 1849 | 25597 | 27405 | 34621 | 38097 |
| LP(人工 RNA) | 3085 | 2000 | 29429 | 29516 | 41849 | 38571 |

２．データの検証

上記３条件でスキャンニングしたアレイ上の遺伝子約７００のデータを取込値としてシグナル値が２０００以上で６００００以下の値のみ使用して、Ｘ軸にＧＰ５０％値、Ｙ軸にＲＬ５０％値をプロットした（図１６、図１７）。

その結果、Ｃｙ３取込値に関しては、相関係数ｒ＝０．９９２、回帰式Ｙ＝０．１０５５Ｘであった（図１６）。Ｃｙ５取込値は、相関係数ｒ＝０．９８６、回帰式Ｙ＝０．０７１５Ｘであった（図１７）。この結果は、Ｃｙ３値であれば、ＲＬ５０％のスキャン条件で取込んだシグナル値は、ＧＰ５０％のスキャン条件で取込んだシグナル値の０．１０５５倍であることを示している。すなわち、スキャン条件が異なっていても同一のスポットのシグナル値は比例関係になっていることを示しており、異なるコントロール毎の取込みが有効であることが示されている。

一方Ｃｙ３とＣｙ５のグラフの位置と傾きを比較すると、両者は一致していない。また両グラフの回帰式も一方が約０．１に対し、もう一方は０．０７と、大きく異なっている。したがってＣｙ３とＣｙ５ではその蛍光強度の傾きが異なるため、即ち低シグナル強度領域ではＣｙ３が高めになり、全体のシグナル強度が増すに従ってＣｙ５が高くなる傾向が推定された。これによりＣｙ３とＣｙ５の比の誤差が、シグナル強度の領域の違いによって、異なることが予想される。

３．各取込みデータの統合（「ＧＰ３万補正」）

そこで上記誤差を解消するため、まず各取込みデータを補正によって統合した。この補正は、Ｃｙ３（またはＣｙ５）の３種類のファイルデータが、同じ傾きの直線であると仮定したものである。具体的にはＣｙ３＿ＧＰ、Ｃｙ５＿ＧＰ、Ｃｙ３＿ＲＬ、Ｃｙ５＿ＲＬ、Ｃｙ３＿ＬＰ、Ｃｙ５＿ＬＰファイルの各ＧＰの７スポットのデータを平均して、各平均値で各ファイルの値を割って、３００００を掛けることによって取込条件が異なるファイル全てのデータが、人工ＲＮＡ　ＧＰの３００００値に合わせた補正データとした。取込時に飽和値（６５５３５）を示しているものは除外し、人工ＲＮＡ　ＧＰ３００００取込の場合は、１０００以下、ＲＬ　３００００取込の場合は、１８００以下を下限値として不採用とした。また、ＬＰ３００００取込の場合は、低発現遺伝子を最も反映できるファイルであるため全て採用した。その結果を以下の表２に示す。

表２

| | Cy3_RL | Cy5_RL | Cy3_GP | Cy5_GP | Cy3_LP | Cy5_LP |
|---|---|---|---|---|---|---|
| 遺伝子名 | GP３万補正 | GP３万補正 | GP３万補正 | GP３万補正 | GP３万補正 | GP３万補正 |
| alpha-1 antitrypsin | | | 6178 | 5247 | 6411 | 5421 |
| alpha-1 antitrypsin | | | 5073 | 4151 | 5511 | 4389 |
| alpha-1 antitrypsin | | | 5574 | 3969 | 5340 | 4161 |
| alpha-1 antitrypsin | | | 6572 | 5598 | 6803 | 5739 |
| NADH dehydrogenase 4 | | | 14972 | 15466 | 15422 | 15903 |
| NADH dehydrogenase 4 | | | 17518 | 18392 | 18111 | 18883 |
| NADH dehydrogenase 4 | | | 14123 | 11063 | 14418 | 10950 |
| NADH dehydrogenase 4 | | | 14360 | 13457 | 15184 | 13581 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 16S rRNA, mitochondrial | 383740 | 504978 | | | | |
| 16S rRNA, mitochondrial | 394577 | 548290 | | | | |
| 16S rRNA, mitochondrial | 376901 | 510288 | | | | |
| 16S rRNA, mitochondrial | 380736 | 514505 | | | | |
| ribosomal protein L41 | 304449 | 331847 | | | | |
| ribosomal protein L41 | 367180 | 370198 | | | | |
| ribosomal protein L41 | 289854 | 311660 | | | | |
| ribosomal protein L41 | 342079 | 352692 | | | | |
| NADH dehydrogenase | | | 17599 | 20905 | 17899 | 21078 |
| NADH dehydrogenase | 22478 | 26446 | 21988 | 26451 | 23477 | 27515 |
| NADH dehydrogenase | | | 13670 | 16052 | 13580 | 15888 |
| NADH dehydrogenase | 16723 | | 18145 | 21945 | 18483 | 22282 |
| Apolipoprotein E | | | 2675 | 1262 | 2421 | 1181 |
| Apolipoprotein E | | | 2539 | 1203 | 2355 | 1241 |
| Apolipoprotein E | | | 2903 | 1239 | 2575 | 1386 |
| Apolipoprotein E | | | 2288 | 1139 | 2979 | 1142 |
| ferritin light chain | 339266 | 388639 | | | | |
| ferritin light chain | 328093 | 362272 | | | | |
| ferritin light chain | 311338 | 343969 | | | | |
| ferritin light chain | 324108 | 317800 | | | | |
| ribosomal protein S12 | 277365 | 369579 | | | | |
| ribosomal protein S12 | 265248 | 274409 | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| ribosomal protein S12 | 284998 | 315013 | | | | |
| ribosomal protein S12 | 285007 | 336248 | | | | |
| ribosomal protein L35 | 96700 | 103715 | | | | |
| ribosomal protein L35 | 105136 | 103630 | | | | |
| ribosomal protein L35 | 91889 | 96876 | | | | |
| ribosomal protein L35 | 99277 | 82909 | | | | |
| ribosomal protein S6 | 329645 | 445141 | | | | |
| ribosomal protein S6 | 350411 | 494950 | | | | |
| ribosomal protein S6 | 274424 | 365659 | | | | |
| ribosomal protein S6 | 306137 | 402541 | | | | |
| kininogen | | | 3295 | 1496 | 3412 | 1693 |
| kininogen | | | 2788 | 1520 | 2630 | 1449 |
| kininogen | | | 3681 | 1692 | 3702 | 1523 |
| kininogen | | | 3170 | 1525 | 3384 | 1379 |
| GAPDH | 244109 | 317853 | | | | |
| GAPDH | 222898 | 285042 | | | | |
| GAPDH | 217638 | 243884 | | | | |
| GAPDH | 225521 | 208129 | | | | |
| thymosin beta-10 | 188698 | 236196 | | | | |
| thymosin beta-10 | 208757 | 247356 | | | | |
| thymosin beta-10 | 208030 | 224231 | | | | |
| thymosin beta-10 | 207958 | 217853 | | | | |
| G6PD | | | 4124 | 2133 | 3738 | 2401 |
| G6PD | | | 4649 | 2543 | 4374 | 2623 |
| G6PD | | | 3803 | 1966 | 3792 | 1985 |

| G6PD | | | 3366 | 1870 | 4073 | 1947 |
|---|---|---|---|---|---|---|
| RL（人工RNA） | 303468 | 416943 | | | | |
| RL（人工RNA） | 281849 | 392157 | | | | |
| RL（人工RNA） | 303841 | 425732 | | | | |
| RL（人工RNA） | 294882 | 427682 | | | | |
| RL（人工RNA） | 278926 | 277953 | | | | |
| RL（人工RNA） | 326722 | 345696 | | | | |
| RL（人工RNA） | 388895 | 427814 | | | | |
| GP（人工RNA） | 33973 | 34642 | 34500 | 35588 | 33585 | 35505 |
| GP（人工RNA） | 26340 | 25708 | 27954 | 27391 | 27539 | 27245 |
| GP（人工RNA） | 24352 | 25669 | 24954 | 25744 | 25408 | 26283 |
| GP（人工RNA） | 26421 | 28225 | 24791 | 28342 | 26777 | 28019 |
| GP（人工RNA） | 27928 | 30953 | 27565 | 28799 | 27656 | 28648 |
| GP（人工RNA） | 29607 | 26103 | 30479 | 26866 | 29434 | 26492 |
| GP（人工RNA） | 41379 | 38701 | 39757 | 37271 | 39601 | 37808 |
| LP(人工RNA) | 22582 | | 24030 | 26799 | 23886 | 26533 |
| LP(人工RNA) | 20072 | | 21944 | 23529 | 21527 | 23263 |
| LP(人工RNA) | 17372 | | 19033 | 16686 | 20285 | 16473 |
| LP(人工RNA) | 18425 | | 19957 | 21596 | 21331 | 21475 |
| LP(人工RNA) | 22369 | | 22128 | 23970 | 22717 | 24229 |
| LP(人工RNA) | 23444 | 24364 | 24342 | 25687 | 24355 | 26934 |
| LP(人工RNA) | 28001 | 26354 | 27986 | 27665 | 29439 | 27269 |
| | Cy3_RL | Cy5_RL | Cy3_GP | Cy5_GP | Cy3_LP | Cy5_LP |

４．各スポットからの遺伝子の仮代表値の決定（棄却検定による例）

（１）ＧＰ３万補正の平均値

さらに上記補正後の各遺伝子スポットの値のどれを採用するか（代表値の選択）について工夫を行った。まず補正後の値についてＣｙ３とＣｙ５それぞれについて平均値を取った。例えば上記表の最後尾のＬＰについては、補正後Ｃｙ３平均は２８４７５（２８００１、２７９８６、２９４３９の平均）となる。

（２）棄却検定

そこでこれらの補正後平均値を用いてスミルノフ・グラッブスの棄却検定を行った。検定の対象は前述のＣｙ３とＣｙ５それぞれについてのＧＰ３万補正後平均値とそれらの比（

Ｃｙ３／Ｃｙ５）について以下の式により行った。
検定統計量Ｔｉ＝｜標本平均値―個々の標本値｜÷√標本分散
有意水準を１０％とすると、ｎ＝４（１遺伝子のスポットの数）の場合、Ｔｉ値が１．４２５以上で棄却、また、ｎ＝７（コントロール遺伝子はスポット数が各７個ある）の場合、Ｔｉ値が１．８２８以上で棄却となる。スポット毎の結果（各行）については、Ｃｙ３、Ｃｙ５、またはＣｙ３／Ｃｙ５のいずれかが棄却となったものについてはデータとして棄却（これ以降の解析から除外する）することとした。結果を以下の表３に示す。

表3

| | | | | 検定統計量Ti | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 遺伝子名 | GP3万補正平均(Cy3) | GP3万補正平均(Cy5) | Cy3/Cy5 | Ti(Cy3) | Ti(Cy5) | Ti(Cy3/Cy5) | Cy3棄却 | Cy5棄却 | Cy3/Cy5比棄却 |
| alpha-1 antitrypsin | 6295 | 5334 | 1.180 | 0.542 | 0.635 | 0.721 | | | |
| alpha-1 antitrypsin | 5292 | 4270 | 1.239 | 0.959 | 0.717 | 0.052 | | | |
| alpha-1 antitrypsin | 5457 | 4065 | 1.342 | 0.713 | 0.978 | 1.396 | | | |
| alpha-1 antitrypsin | 6687 | 5668 | 1.180 | 1.130 | 1.060 | 0.726 | | | |
| NADH dehydrogenase 4 | 15197 | 15685 | 0.969 | 0.200 | 0.300 | 0.694 | | | |
| NADH dehydrogenase 4 | 17814 | 18638 | 0.956 | 1.456 | 1.211 | 0.777 | 棄却 | | |
| NADH dehydrogenase 4 | 14270 | 11006 | 1.297 | 0.787 | 1.143 | 1.381 | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| NADH dehydrogenase 4 | 14772 | 13519 | 1.093 | 0.469 | 0.368 | 0.090 | | | |
| 16S rRNA, mitochondrial | 383740 | 504978 | 0.760 | 0.033 | 0.743 | 1.238 | | | |
| 16S rRNA, mitochondrial | 394577 | 548290 | 0.720 | 1.394 | 1.470 | 1.209 | | 棄却 | |
| 16S rRNA, mitochondrial | 376901 | 510288 | 0.739 | 0.933 | 0.471 | 0.057 | | | |
| 16S rRNA, mitochondrial | 380736 | 514505 | 0.740 | 0.428 | 0.256 | 0.028 | | | |
| ribosomal protein L41 | 304449 | 331847 | 0.917 | 0.608 | 0.384 | 1.009 | | | |
| ribosomal protein L41 | 367180 | 370198 | 0.992 | 1.172 | 1.127 | 1.144 | | | |
| ribosomal protein L41 | 289854 | 311660 | 0.930 | 1.023 | 1.180 | 0.644 | | | |
| ribosomal protein | 342079 | 352692 | 0.970 | 0.459 | 0.437 | 0.509 | | | |

| L41 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| NADH dehydrogenase | 17749 | 20992 | 0.846 | 0.055 | 0.108 | 0.387 | | | |
| NADH dehydrogenase | 22647 | 26804 | 0.845 | 1.273 | 1.199 | 0.360 | | | |
| NADH dehydrogenase | 13625 | 15970 | 0.853 | 1.173 | 1.237 | 0.732 | | | |
| NADH dehydrogenase | 17784 | 22114 | 0.804 | 0.045 | 0.145 | 1.478 | | | 棄却 |
| Apolipoprotein E | 2548 | 1221 | 2.086 | 0.353 | 0.041 | 0.264 | | | |
| Apolipoprotein E | 2447 | 1222 | 2.002 | 1.166 | 0.027 | 0.906 | | | |
| Apolipoprotein E | 2739 | 1312 | 2.087 | 1.185 | 1.258 | 0.260 | | | |
| Apolipoprotein E | 2633 | 1140 | 2.309 | 0.335 | 1.190 | 1.429 | | | 棄却 |
| ferritin light chain | 339266 | 388639 | 0.873 | 1.177 | 1.187 | 0.821 | | | |
| ferritin light | 328093 | 362272 | 0.906 | 0.207 | 0.305 | 0.314 | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| chain | | | | | | | | | |
| ferritin light chain | 311338 | 343969 | 0.905 | 1.246 | 0.308 | 0.322 | | | |
| ferritin light chain | 324108 | 317800 | 1.020 | 0.138 | 1.184 | 1.457 | | | 棄却 |
| ribosomal protein S12 | 277365 | 369579 | 0.750 | 0.085 | 1.148 | 1.273 | | | |
| ribosomal protein S12 | 265248 | 274409 | 0.967 | 1.384 | 1.239 | 1.081 | | | |
| ribosomal protein S12 | 284998 | 315013 | 0.905 | 0.734 | 0.221 | 0.407 | | | |
| ribosomal protein S12 | 285007 | 336248 | 0.848 | 0.735 | 0.312 | 0.215 | | | |
| ribosomal protein L35 | 96700 | 103715 | 0.932 | 0.281 | 0.708 | 0.748 | | | |
| ribosomal protein L35 | 105136 | 103630 | 1.015 | 1.248 | 0.700 | 0.071 | | | |
| ribosomal | 91889 | 96876 | 0.949 | 1.153 | 0.010 | 0.615 | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| protein L35 | | | | | | | | | |
| ribosomal protein L35 | 99277 | 82909 | 1.197 | 0.186 | 1.417 | 1.434 | | | 棄却 |
| ribosomal protein S6 | 329645 | 445141 | 0.741 | 0.444 | 0.324 | 0.029 | | | |
| ribosomal protein S6 | 350411 | 494950 | 0.708 | 1.081 | 1.219 | 1.401 | | | |
| ribosomal protein S6 | 274424 | 365659 | 0.750 | 1.248 | 1.103 | 0.466 | | | |
| ribosomal protein S6 | 306137 | 402541 | 0.761 | 0.276 | 0.440 | 0.906 | | | |
| kininogen | 3353 | 1594 | 2.104 | 0.235 | 0.767 | 0.079 | | | |
| kininogen | 2709 | 1485 | 1.825 | 1.345 | 0.639 | 1.382 | | | |
| kininogen | 3692 | 1607 | 2.297 | 1.064 | 0.935 | 0.824 | | | |
| kininogen | 3277 | 1452 | 2.257 | 0.046 | 1.063 | 0.637 | | | |
| GAPDH | 244109 | 317853 | 0.768 | 1.438 | 1.131 | 0.779 | 棄却 | | |
| GAPDH | 222898 | 285042 | 0.782 | 0.403 | 0.445 | 0.683 | | | |
| GAPDH | 217638 | 243884 | 0.892 | 0.860 | 0.415 | 0.075 | | | |
| GAPDH | 225521 | 208129 | 1.084 | 0.175 | 1.162 | 1.386 | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| thymosin beta-10 | 188698 | 236196 | 0.799 | 1.499 | 0.366 | 1.139 | 棄却 | | |
| thymosin beta-10 | 208757 | 247356 | 0.844 | 0.552 | 1.220 | 0.516 | | | |
| thymosin beta-10 | 208030 | 224231 | 0.928 | 0.477 | 0.549 | 0.642 | | | |
| thymosin beta-10 | 207958 | 217853 | 0.955 | 0.470 | 1.037 | 1.013 | | | |
| G6PD | 3931 | 2267 | 1.734 | 0.165 | 0.270 | 0.917 | | | |
| G6PD | 4512 | 2583 | 1.747 | 1.455 | 1.295 | 0.804 | 棄却 | | |
| G6PD | 3797 | 1976 | 1.922 | 0.537 | 0.674 | 0.741 | | | |
| G6PD | 3720 | 1908 | 1.949 | 0.753 | 0.892 | 0.981 | | | |
| RL (人工RNA) | 303468 | 416943 | 0.728 | 0.205 | 0.515 | 0.664 | | | |
| RL (人工RNA) | 281849 | 392157 | 0.719 | 0.778 | 0.078 | 0.733 | | | |
| RL (人工RNA) | 303841 | 425732 | 0.714 | 0.196 | 0.670 | 0.771 | | | |
| RL (人工RNA) | 294882 | 427682 | 0.689 | 0.433 | 0.704 | 0.955 | | | |
| RL (人工RNA) | 278926 | 277953 | 1.003 | 0.855 | 1.933 | 1.428 | | | |
| RL (人工RNA) | 326722 | 345696 | 0.945 | 0.410 | 0.740 | 0.985 | | | |
| RL (人工 | 388895 | 427814 | 0.909 | 2.056 | 0.706 | 0.711 | 棄 | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| RNA) | | | | | | | 却 | | |
| GP（人工RNA) | 34019 | 35245 | 0.965 | 0.743 | 1.112 | 0.465 | | | |
| GP（人工RNA) | 27278 | 26781 | 1.019 | 0.503 | 0.683 | 0.261 | | | |
| GP（人工RNA) | 24905 | 25899 | 0.962 | 0.941 | 0.870 | 0.514 | | | |
| GP（人工RNA) | 25996 | 28195 | 0.922 | 0.740 | 0.383 | 1.053 | | | |
| GP（人工RNA) | 27716 | 29466 | 0.941 | 0.422 | 0.113 | 0.800 | | | |
| GP（人工RNA) | 29840 | 26487 | 1.127 | 0.030 | 0.745 | 1.731 | | | |
| GP（人工RNA) | 40246 | 37926 | 1.061 | 1.893 | 1.681 | 0.840 | | | |
| LP(人工RNA) | 23499 | 26666 | 0.881 | 0.274 | 0.848 | 0.919 | | | |
| LP(人工RNA) | 21181 | 23396 | 0.905 | 0.456 | 0.049 | 0.660 | | | |
| LP(人工RNA) | 18896 | 16579 | 1.140 | 1.176 | 1.920 | 1.854 | | | |
| LP(人工RNA) | 19904 | 21535 | 0.924 | 0.859 | 0.560 | 0.457 | | | |
| LP(人工RNA) | 22405 | 24099 | 0.930 | 0.071 | 0.144 | 0.399 | | | |
| LP(人工RNA) | 24047 | 25662 | 0.937 | 0.446 | 0.572 | 0.320 | | | |
| RNA) | | | | | | | | | |
| LP(人工RNA) | 28475 | 27096 | 1.051 | 1.842 | 0.966 | 0.901 | | | |

（３）仮代表値の決定

棄却された値を除いた後、平均値または中央値を求めて遺伝子の仮代表値とすることができる。ここでは、平均値の結果を以下の表４に示す。

表４

| 遺伝子名 | Cy3_GP棄却後GP3万補正平均 | Cy5_GP棄却後GP3万補正平均 | 遺伝子仮代表値Cy3 | 遺伝子仮代表値Cy5 |
|---|---|---|---|---|
| alpha-1 antitrypsin | 6295 | 5334 | 5933 | 4834 |
| alpha-1 antitrypsin | 5292 | 4270 | | |
| alpha-1 antitrypsin | 5457 | 4065 | | |
| alpha-1 antitrypsin | 6687 | 5668 | | |
| NADH dehydrogenase 4 | 15197 | 15685 | 14746 | 13403 |
| NADH dehydrogenase 4 | | | | |
| NADH dehydrogenase 4 | 14270 | 11006 | | |
| NADH | 14772 | 13519 | | |

| | | | | |
|---|---|---|---|---|
| dehydrogenase 4 | | | | |
| 16S rRNA, mitochondrial | 383740 | 504978 | 380459 | 509923 |
| 16S rRNA, mitochondrial | | | | |
| 16S rRNA, mitochondrial | 376901 | 510288 | | |
| 16S rRNA, mitochondrial | 380736 | 514505 | | |
| ribosomal protein L41 | 304449 | 331847 | 325891 | 341599 |
| ribosomal protein L41 | 367180 | 370198 | | |
| ribosomal protein L41 | 289854 | 311660 | | |
| ribosomal protein L41 | 342079 | 352692 | | |
| NADH dehydrogenase | 17749 | 20992 | 18007 | 21255 |
| NADH dehydrogenase | 22647 | 26804 | | |
| NADH dehydrogenase | 13625 | 15970 | | |
| NADH dehydrogenase | | | | |

| | | | | |
|---|---|---|---|---|
| Apolipoprotein E | 2548 | 1221 | 2578 | 1252 |
| Apolipoprotein E | 2447 | 1222 | | |
| Apolipoprotein E | 2739 | 1312 | | |
| Apolipoprotein E | | | | |
| ferritin light chain | 339266 | 388639 | 326232 | 364960 |
| ferritin light chain | 328093 | 362272 | | |
| ferritin light chain | 311338 | 343969 | | |
| ferritin light chain | | | | |
| ribosomal protein S12 | 277365 | 369579 | 278154 | 323812 |
| ribosomal protein S12 | 265248 | 274409 | | |
| ribosomal protein S12 | 284998 | 315013 | | |
| ribosomal protein S12 | 285007 | 336248 | | |
| ribosomal protein L35 | 96700 | 103715 | 97908 | 101407 |

| | | | | |
|---|---|---|---|---|
| ribosomal protein L35 | 105136 | 103630 | | |
| ribosomal protein L35 | 91889 | 96876 | | |
| ribosomal protein L35 | | | | |
| ribosomal protein S6 | 329645 | 445141 | 315154 | 427073 |
| ribosomal protein S6 | 350411 | 494950 | | |
| ribosomal protein S6 | 274424 | 365659 | | |
| ribosomal protein S6 | 306137 | 402541 | | |
| kininogen | 3353 | 1594 | 3258 | 1535 |
| kininogen | 2709 | 1485 | | |
| kininogen | 3692 | 1607 | | |
| kininogen | 3277 | 1452 | | |
| GAPDH | | | 222019 | 245685 |
| GAPDH | 222898 | 285042 | | |
| GAPDH | 217638 | 243884 | | |
| GAPDH | 225521 | 208129 | | |
| thymosin beta-10 | | | 208248 | 229814 |
| thymosin beta-10 | 208757 | 247356 | | |

| | | | | |
|---|---|---|---|---|
| thymosin beta-10 | 208030 | 224231 | | |
| thymosin beta-10 | 207958 | 217853 | | |
| G6PD | 3931 | 2267 | 3816 | 2050 |
| G6PD | | | | |
| G6PD | 3797 | 1976 | | |
| G6PD | 3720 | 1908 | | |
| RL (人工RNA) | 303468 | 416943 | 298281 | 381027 |
| RL (人工RNA) | 281849 | 392157 | | |
| RL (人工RNA) | 303841 | 425732 | | |
| RL (人工RNA) | 294882 | 427682 | | |
| RL (人工RNA) | 278926 | 277953 | | |
| RL (人工RNA) | 326722 | 345696 | | |
| RL (人工RNA) | | | | |
| GP (人工RNA) | 34019 | 35245 | 30000 | 30000 |
| GP (人工RNA) | 27278 | 26781 | | |
| GP (人工RNA) | 24905 | 25899 | | |
| GP (人工RNA) | 25996 | 28195 | | |
| GP (人工RNA) | 27716 | 29466 | | |
| GP (人工RNA) | 29840 | 26487 | | |
| GP (人工RNA) | 40246 | 37926 | | |
| LP(人工RNA) | 23499 | 26666 | 22630 | 23576 |
| LP(人工RNA) | 21181 | 23396 | | |
| LP(人工RNA) | 18896 | 16579 | | |
| LP(人工RNA) | 19904 | 21535 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| LP(人工RNA) | 22405 | 24099 | | | | | |
| LP(人工RNA) | 24047 | 25662 | | | | | |
| LP(人工RNA) | 28475 | 27096 | | | | | |

５．各スポットからの遺伝子の仮代表値の決定（ユークリッド距離を用いた外れ値の除外）

前述の棄却検定に代わって、ユークリッド距離を用いて外れ値を除外する方法でも仮代表値を決定することができる。

まず前述の棄却検定と同様に、ＧＰ３万補正後の平均値を求めた。さらにその値から下記の計算式によりユークリッド距離を求め、それぞれの距離値を比較し、外れ値を除外した。

計算式：

√（標本Ｃｙ３平均値―個々のＣｙ３標本値）$^2$ ＋（標本Ｃｙ３平均値―個々のＣｙ３標本値）$^2$

棄却（除外）の選択方法は、最も近いものを４個中２個選択するなど色々な方法で行うことができるが、本実施例では同じ遺伝子スポット（４個または７個）中で最も大きく外れたもの１個を棄却とした。

さらに外れ値を除外した値の平均値をとり、各遺伝子の仮代表値とした。これらの仮代表値は、遺伝子によっては前述の方法による仮代表値と異なる。その結果を以下の表５に示す。

表５

| 遺伝子名 | Cy3 GP3万補 | Cy5 GP3万補 | Cy3/Cy5 | ユークリッド | Cy3 | 遺伝子仮代表値 | 遺伝子仮代表値 |
|---|---|---|---|---|---|---|---|

| | 正平均 | 正平均 | | 距離 | 棄却 | Cy3 | Cy5 |
|---|---|---|---|---|---|---|---|
| alpha-1 antitrypsin | 6295 | 5334 | 1.180 | 617 | | 5681 | 4556 |
| alpha-1 antitrypsin | 5292 | 4270 | 1.239 | 854 | | | |
| alpha-1 antitrypsin | 5457 | 4065 | 1.342 | 905 | | | |
| alpha-1 antitrypsin | 6687 | 5668 | 1.180 | 1125 | 棄却 | | |
| NADH dehydrogenase 4 | 15197 | 15685 | 0.969 | 1023 | | 14746 | 13403 |
| NADH dehydrogenase 4 | 17814 | 18638 | 0.956 | 4550 | 棄却 | | |
| NADH dehydrogenase 4 | 14270 | 11006 | 1.297 | 3909 | | | |
| NADH dehydrogenase 4 | 14772 | 13519 | 1.093 | 1404 | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16S rRNA, mitochondrial | 383740 | 504978 | 0.760 | 14540 | | 380459 | 509923 |
| 16S rRNA, mitochondrial | 394577 | 548290 | 0.720 | 30661 | 棄却 | | |
| 16S rRNA, mitochondrial | 376901 | 510288 | 0.739 | 11635 | | | |
| 16S rRNA, mitochondrial | 380736 | 514505 | 0.740 | 5974 | | | |
| ribosomal protein L41 | 304449 | 331847 | 0.917 | 23556 | | 312127 | 332066 |
| ribosomal protein L41 | 367180 | 370198 | 0.992 | 50227 | 棄却 | | |
| ribosomal protein L41 | 289854 | 311660 | 0.930 | 46851 | | | |
| ribosomal protein L41 | 342079 | 352692 | 0.970 | 19625 | | | |
| NADH dehydrogen | 17749 | 20992 | 0.846 | 519 | | 16386 | 19692 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ase | | | | | | | |
| NADH dehydrogenase | 22647 | 26804 | 0.845 | 7107 | 棄却 | | |
| NADH dehydrogenase | 13625 | 15970 | 0.853 | 6997 | | | |
| NADH dehydrogenase | 17784 | 22114 | 0.804 | 665 | | | |
| Apolipoprotein E | 2548 | 1221 | 2.086 | 44 | | 2543 | 1195 |
| Apolipoprotein E | 2447 | 1222 | 2.002 | 145 | | | |
| Apolipoprotein E | 2739 | 1312 | 2.087 | 172 | 棄却 | | |
| Apolipoprotein E | 2633 | 1140 | 2.309 | 93 | | | |
| ferritin light chain | 339266 | 388639 | 0.873 | 37974 | 棄却 | 321179 | 341347 |
| ferritin light chain | 328093 | 362272 | 0.906 | 9411 | | | |
| ferritin light | 311338 | 343969 | 0.905 | 17058 | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| chain | | | | | | | |
| ferritin light chain | 324108 | 317800 | 1.020 | 35406 | | | |
| ribosomal protein S12 | 277365 | 369579 | 0.750 | 45773 | | 282457 | 340280 |
| ribosomal protein S12 | 265248 | 274409 | 0.967 | 51062 | 棄却 | | |
| ribosomal protein S12 | 284998 | 315013 | 0.905 | 11147 | | | |
| ribosomal protein S12 | 285007 | 336248 | 0.848 | 14199 | | | |
| ribosomal protein L35 | 96700 | 103715 | 0.932 | 7104 | | 97908 | 101407 |
| ribosomal protein L35 | 105136 | 103630 | 1.015 | 9710 | | | |
| ribosomal protein L35 | 91889 | 96876 | 0.949 | 6362 | | | |
| ribosomal | 99277 | 82909 | 1.197 | 13912 | 棄 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| protein L35 | | | | | 却 | | |
| ribosomal protein S6 | 329645 | 445141 | 0.741 | 23161 | | 303402 | 404447 |
| ribosomal protein S6 | 350411 | 494950 | 0.708 | 76488 | 棄却 | | |
| ribosomal protein S6 | 274424 | 365659 | 0.750 | 73693 | | | |
| ribosomal protein S6 | 306137 | 402541 | 0.761 | 26137 | | | |
| kininogen | 3353 | 1594 | 2.104 | 113 | | 3441 | 1551 |
| kininogen | 2709 | 1485 | 1.825 | 551 | 棄却 | | |
| kininogen | 3692 | 1607 | 2.297 | 440 | | | |
| kininogen | 3277 | 1452 | 2.257 | 85 | | | |
| GAPDH | 244109 | 317853 | 0.768 | 56605 | 棄却 | 222019 | 245685 |
| GAPDH | 222898 | 285042 | 0.782 | 21815 | | | |
| GAPDH | 217638 | 243884 | 0.892 | 22177 | | | |
| GAPDH | 225521 | 208129 | 1.084 | 55635 | | | |
| thymosin beta-10 | 188698 | 236196 | 0.799 | 15424 | | 201562 | 226093 |
| thymosin beta-10 | 208757 | 247356 | 0.844 | 16835 | 棄却 | | |
| thymosin beta-10 | 208030 | 224231 | 0.928 | 8563 | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| thymosin beta-10 | 207958 | 217853 | 0.955 | 14314 | | | |
| G6PD | 3931 | 2267 | 1.734 | 102 | | 3816 | 2050 |
| G6PD | 4512 | 2583 | 1.747 | 657 | 棄却 | | |
| G6PD | 3797 | 1976 | 1.922 | 283 | | | |
| G6PD | 3720 | 1908 | 1.949 | 386 | | | |
| RL（人工RNA） | 303468 | 416943 | 0.728 | 30244 | | 316610 | 406004 |
| RL（人工RNA） | 281849 | 392157 | 0.719 | 29712 | | | |
| RL（人工RNA） | 303841 | 425732 | 0.714 | 38732 | | | |
| RL（人工RNA） | 294882 | 427682 | 0.689 | 30244 | | | |
| RL（人工RNA） | 278926 | 277953 | 1.003 | 114412 | 棄却 | | |
| RL（人工RNA） | 326722 | 345696 | 0.945 | 44782 | | | |
| RL（人工RNA） | 388895 | 427814 | 0.909 | 87411 | | | |
| GP（人工RNA） | 34019 | 35245 | 0.965 | 6608 | | 28292 | 28679 |
| GP（人工RNA） | 27278 | 26781 | 1.019 | 4216 | | | |
| GP（人工 | 24905 | 25899 | 0.962 | 6541 | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| RNA) | | | | | | | |
| GP（人工RNA） | 25996 | 28195 | 0.922 | 6608 | | | |
| GP（人工RNA） | 27716 | 29466 | 0.941 | 2345 | | | |
| GP（人工RNA） | 29840 | 26487 | 1.127 | 3517 | | | |
| GP（人工RNA） | 40246 | 37926 | 1.061 | 12954 | 棄却 | | |
| LP(人工RNA) | 23499 | 26666 | 0.881 | 3210 | | 23252 | 24742 |
| LP(人工RNA) | 21181 | 23396 | 0.905 | 1460 | | | |
| LP(人工RNA) | 18896 | 16579 | 1.140 | 7931 | 棄却 | | |
| LP(人工RNA) | 19904 | 21535 | 0.924 | 3210 | | | |
| LP(人工RNA) | 22405 | 24099 | 0.930 | 569 | | | |
| LP(人工RNA) | 24047 | 25662 | 0.937 | 2521 | | | |
| LP(人工RNA) | 28475 | 27096 | 1.051 | 6823 | | | |

６．Ｃｙ３とＣｙ５のゆがみの補正：代表値の決定

前述のＧＰ３万補正によって、グラフ上の右上と左下にあったグラフを中央のグラフに合わせこみ、さらに棄却検定等により、外れ値を除外した。この作業により、測定値のほとんどが１本の直線上にのることとなったが、こんどはこの直線のゆがみが問題となる（図１８）。

図１８ではＣｙ３側ＲＮＡ、Ｃｙ５側ＲＮＡは同じＲＮＡである。従ってＣｙ３とＣｙ５が全く同じ挙動を示すならばグラフ中央の直線と同じになるはずである。ところが低発現域ではＣｙ３がＣｙ５より強い。これは先のＧＰ３万補正を行っても現れるため、Ｃｙ３とＣｙ５の誤差ではなく、支持体であるスライドグラスがわずかにＣｙ３と近い波長の蛍光を発していることが原因と推定された。

実際、このゆがみは、相加的なゆがみであることを見つけ出している。例として、無味数での計算結果を示す。図１９において擬似Ｃｙ５値は、１００００から２で割った等比数

列で、擬似Ｃｙ３値は、擬似Ｃｙ５値を１．２倍して３０を加えたものであるが、マイクロアレイのＣｙ３Ｃｙ５値のゆがみと酷似している。
つまり、この図１９では、Ｃｙ３に３０加えただけで、上記のようなゆがみが生じさせることができる。
そこで、幾つかのスポットを用いて相関関係より回帰式を求めて、片側の値（例えばＣｙ５値）を補正することによってこのゆがみは解消されるはずである。例えば上記では、回帰直線の式は（擬似Ｃｙ３）＝１．２ｘ（擬似Ｃｙ５）＋３０となるため、擬似Ｃｙ５値を１．２倍して３０を加えれば直線に変換できる。また、擬似Ｃｙ３値から３０を引いて１．２で割れば直線にできる。すなわち、回帰直線の式を用いることによって１：１の関係に補正できることになる。
そこで図１８で示した実際のデータを用いて回帰式の補正ができるか検証した。補正用遺伝子として約２００遺伝子を選択して（後述「７．コントロールＲＮＡ（レファレンスＲＮＡ）の補正）、その回帰式（擬似Ｃｙ５）＝１．２ｘ（擬似Ｃｙ３）－２２７２を用いて補正した。その結果を図２０に示す。
この補正によって、直線関係が得られない低発現部分に関しては、これまで測定感度以下として扱われてきたが、評価対象にすることができることが示された。そこで、この補正で得られたＣｙ３値とＣｙ５値を代表値として決定する。
７．コントロールＲＮＡ（レファレンスＲＮＡ）の補正
前述のようにＤＮＡマイクロアレイではアレイ間差をなくすため、検体ＲＮＡに対して対照ＲＮＡを同時に測定し、比として結果を得ることが多い。具体的にはＣｙ３またはＣｙ５の一方を例えば培養細胞からの抽出ＲＮＡとし、もう一方を検体ＲＮＡとして、それぞれをＣｙ３またはＣｙ５で標識して、最終的にはアレイの各スポットについてＣｙ３とＣｙ５との比で求めるというものである。
この場合実験開始時に両者のＲＮＡ量を揃えても、Ｃｙ３Ｃｙ５の標識効率が異なるため、あるいは標識工程等での量の減少が想定されるため、単純にマイクロアレイ上の測定値を比較することはできない。
一般にはハウスキーピング遺伝子を用いて補正を行うが、培養細胞では必ずしもハウスキーピング遺伝子がハウスキーピングに働いているとは限らず、選択が難しい。
（１）補正用遺伝子の選択方法
そこでＣｙ５標識したコントロール（培養細胞）ＲＮＡとＣｙ３標識した検体（ヒト肝組織）ＲＮＡを同一のマイクロアレイ上でハイブリダイゼーションさせ、ほぼ１対１の関係にある遺伝子を選択した（図２１、図２２）。
図中では選択した補正用遺伝子を■、その他の遺伝子を○で示した。検体２例分を示す。
このような選択方法であれば、コントロールＲＮＡと検体ＲＮＡが類似していなくとも補正用遺伝子の選択はできる。
なお、無処置と薬剤処置、病態Ａと病態Ｂ、薬剤応答性なしと薬剤応答性あり等の比較を望む場合は以下のように行う。たとえば、無処置と薬剤処置を各々１０例がある場合は、まずそれぞれの比較したい群に分け、各々１０例のデータを取得してＧＰ３万補正を行い、各々のＣｙ３／Ｃｙ５比を求める。この合計２０例の全ての検体でＣｙ３／Ｃｙ５比が２～０．５の範囲内の遺伝子セットを割り出す。この時、ばらつきの指標であるＣＶ値を求めて、２０％以内の遺伝子セットにすることもできる。さらに好ましくは、Ｃｙ３／Ｃｙ５比が１．５～１／１．５の範囲内かつＣＶ値１０％以内の遺伝子セットである。ここで選択された遺伝子セットを補正用遺伝子として使用する。
この遺伝子セット（補正用遺伝子）を利用して回帰式の補正を行う。バックグランドに相当する溶媒のみのスポットやブランク用スポットのＣｙ３／Ｃｙ５値が、選択した遺伝子セットを用いた回帰式補正で、２～０．５の範囲内になるか判断する。もし、バックグランドに相当する溶媒のみのスポットやブランク用スポットのＣｙ３／Ｃｙ５値がこの範囲より外れる場合、補正用遺伝子選択時に使用した例数を増加する。この過程を繰り返すことによって、最適な補正用遺伝子の選択を行う。
（２）補正結果

この補正用遺伝子を用いてデータ全体を補正した。つまり、一般にはノーマライズはメンテナンス遺伝子（内部標準）か、人工的に外部から入れたスパイク（コントロール）のいずれかで行うが、本実施例ではＣｙ３内またはＣｙ５内のファイルの統合をスパイクで行い、Ｃｙ３とＣｙ５間のノーマライズを上記によって選択した遺伝子で行った。補正した結果を図２３、図２４に示す。

８．補正の有効性の検証

これまで述べた、異なる複数のコントロールＲＮＡに基づくスキャニングとその統合及び補正をしたデータを用いて解析を行った。Ｃ型肝炎患者の肝組織から抽出したＲＮＡについて、ＤＮＡマイクロアレイ（自家製）を用いてハイブリダイゼーションを行い、患者１０名のデータを取得した。このデータについて、前記実施例と同様に、異なったレーザーパワーで取込んだ値を、ノーマライズ用遺伝子（ＲＬ、ＧＰ、ＬＰ）の値を用いて統合した。その値をＧＰ３万補正して三井情報開発株式会社（会社名）のＧｅｎｏｍｉｃＰｒｏｆｉｌｅｒソフトウエア（商品名）を用いてＥｕｃｌｉｄｅａｎ　Ｗａｒｄ法で階層クラスタリングを行ったところ、検体は大きく２つのクラスターに分かれ、片方のクラスターはさらに２つに分かれた（図２５）。図中のサンプル番号についているＮ及びＣは、臨床的に診断されたもので、Ｃは薬剤（インターフェロン療法）の効果を示すタイプを示し、Ｎは薬剤の効果が無いタイプを示している。クラスターが３つに分かれたことにより、ＧＰ３万補正だけでは薬剤が有効か無効か判断のつかない群が現れ、臨床診断との不一致部分が残った。

そこで、ＧＰ３万補正の後に、スミルノフ・グラッブスの棄却検定を用いたはずれ値の棄却（上述４（１））、残りの平均値による代表値の決定（上述４（２））、Ｃｙ３・Ｃｙ５のゆがみ補正（上述６）を上述７（１）で決定した補正用遺伝子（Ｉｎｔｅｒｎａｌ　ｃｏｎｔｒｏｌ）を用いたノーマライズを行い、再度同じ解析方法で検証した。

その結果、先の１０検体は大きく２群のクラスターに分かれ、臨床診断上の薬剤無効群と効果群と一致した（図２６）。すなわち、マイクロアレイデータの取得方法として、異なった取り込み条件でデータを取得することによって１条件の取り込みでは得られない、高シグナルから低シグナルまでの精度の高いデータが取得できること、これら異なった条件で取り込んだデータを一つの条件で取り込んだデータとして取り扱いができるように統合できること、棄却過程によって信頼性の高いデータにできること、補正用遺伝子を用いて回帰式を利用した補正ができること、これら補正を行ったマイクロアレイデータを使用することにより判定解析ができることを示している。

産業上の利用可能性

以上のように、ＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体は、従来は一種類とされていたスパイクＲＮＡやハウスキーピング遺伝子などのコントロール用遺伝子に複数の遺伝子ＲＮＡを使用することにより、ＤＮＡマイクロアレイによる測定精度や再現性の向上を可能とし、測定レンジを拡大させ、また、ＲＮＡの増幅や標識反応、アレイ上での結合（ハイブリダイゼーション）反応、およびシグナル取込み等の一連の過程において現れる様々なばらつきや誤差を、補正によって取り除くことにより信頼性の高いデータを取得し、さらにそれらの補正方法を一般生物学研究者が容易に理解しうる構成とすることで、ユーザーが自由に操作・活用でき、同時に作業の煩雑さを軽減し効率化を図り、スキャニングからデータ解析に至る完全自動化を可能とする。

これにより、本発明にかかるＤＮＡマイクロアレイデータ処理方法、ＤＮＡマイクロアレイデータ処理システム、ＤＮＡマイクロアレイデータ処理装置、プログラム、および、記録媒体は、ＤＮＡ配列などの解析を行うバイオインフォマティクス分野において極めて有用である。すなわち、本発明は、産業上多くの分野、特に医薬品、食品、化粧品、医療、遺伝子発現解析等の分野で広く実施することができ、極めて有用である。

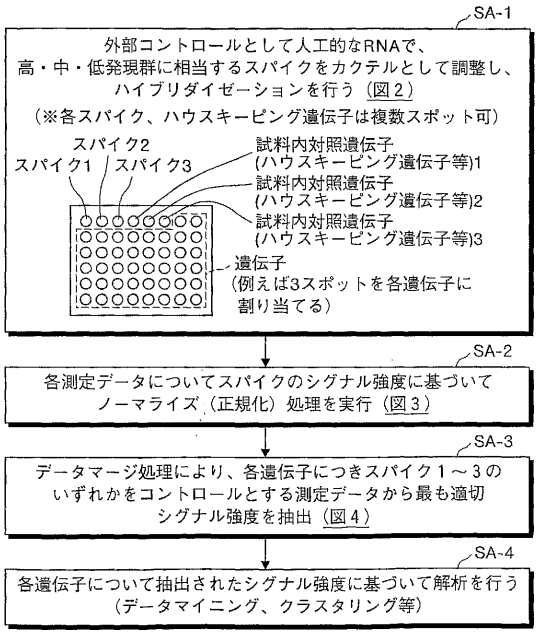
【図面の簡単な説明】

第１図は、本発明の基本原理を示す原理構成図であり、第２図は、複数のスパイクＲＮＡを用いた二蛍光標識法によるスパイク実験の概念を説明する概念図であり、第３図は、ス

キャナ装置を用いてＤＮＡマイクロアレイの発光支持体をスキャニングを行い、各スパイクＲＮＡのシグナル強度に基づいてノーマライゼーションを行う概念を説明するフロー図であり、第４図は、各測定対象の遺伝子ＲＮＡについて最適なシグナル強度を選択する場合の概念を説明するフロー図であり、第５図は、本発明が適用される本システムの構成の一例を示すブロック図であり、第６図は、図５に示すＤＮＡマイクロアレイデータ処理装置１００の機能ブロックの一例を示す図であり、第７図は、図５に示すＤＮＡマイクロアレイスキャニングシステムの機能ブロックの一例を示す図であり、第８図は、本実施形態における本システムのスキャニング処理の一例を示すフローチャートであり、第９図は、本実施形態における本システムのスキャニング画像取得処理の一例を示すフローチャートであり、第１０図は、本実施形態における本システムのノーマライズ処理の一例を示すフローチャートであり、第１１図は、本実施形態における本システムのシグナル強度選択処理の一例を示すフローチャートであり、第１２図は、ノーマライズ部１０２ｂの構成の一例を示すブロック図であり、第１３図は、代表値決定部１０２ｊの構成の一例を示すブロック図であり、第１４図は、本実施形態における本システムのノーマライズ処理の別の実施形態の一例を示すフローチャートであり、第１５図は、スキャナ装置２００によるスキャニング条件の設定の一例を示すフローチャートであり、第１６図は、３条件でスキャンニングしたアレイ上の遺伝子約７００のデータを取込値としてシグナル値が２０００以上で６００００以下の値のみ使用して、Ｘ軸にＧＰ５０％値、Ｙ軸にＲＬ５０％値をプロットした図であり、第１７図は、３条件でスキャンニングしたアレイ上の遺伝子約７００のデータを取込値としてシグナル値が２０００以上で６００００以下の値のみ使用して、Ｘ軸にＧＰ５０％値、Ｙ軸にＲＬ５０％値をプロットした図であり、第１８図は、ＧＰ３万補正によって、グラフ上の右上と左下にあったグラフを中央のグラフに合わせこみ、さらに棄却検定等により、外れ値を除外した値をプロットした図であり、第１９図は、擬似Ｃｙ５値が１００００から２で割った等比数列であり、擬似Ｃｙ３値が擬似Ｃｙ５値を１．２倍して３０を加えた値をプロットした図であり、第２０図は、第１８図の結果を補正用遺伝子として約２００遺伝子を選択してその回帰式（擬似Ｃｙ５）＝１．２ｘ（擬似Ｃｙ３）－２２７２を用いて補正した結果を示す図である。第２１図は、Ｃｙ５標識したコントロール（培養細胞）ＲＮＡとＣｙ３標識した検体（ヒト肝組織）ＲＮＡを同一のマイクロアレイ上でハイブリダイゼーションさせ、ほぼ１対１の関係にある遺伝子を選択した図であり、１対１の関係にある遺伝子を■で表し、それ以外のものを○で示している。第２２図は、Ｃｙ５標識したコントロール（培養細胞）ＲＮＡとＣｙ３標識した検体（ヒト肝組織）ＲＮＡを同一のマイクロアレイ上でハイブリダイゼーションさせ、ほぼ１対１の関係にある遺伝子を選択した図であり、１対１の関係にある遺伝子を■で表し、それ以外のものを○で示している。第２３図は、第２１図の補正用遺伝子を使用して回帰式を用いた補正を行った結果を示す図であり、１対１の関係にある遺伝子を■で表し、それ以外のものを○で示している。第２４図は、第２２図の補正用遺伝子を使用して回帰式を用いた補正を行った結果を示す図であり、１対１の関係にある遺伝子を■で表し、それ以外のものを○で示している。第２５図は、Ｅｕｃｌｉｄｅａｎ　Ｗａｒｄ法で階層クラスタリングを行った結果を示す図であり、第２６図は、補正後にＥｕｃｌｉｄｅａｎ　Ｗａｒｄ法で階層クラスタリングを行った結果を示す図である。

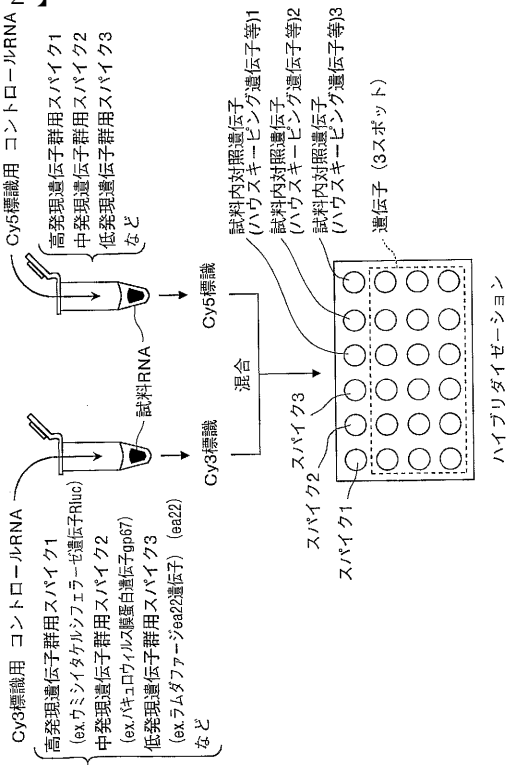
【図１】

第１図



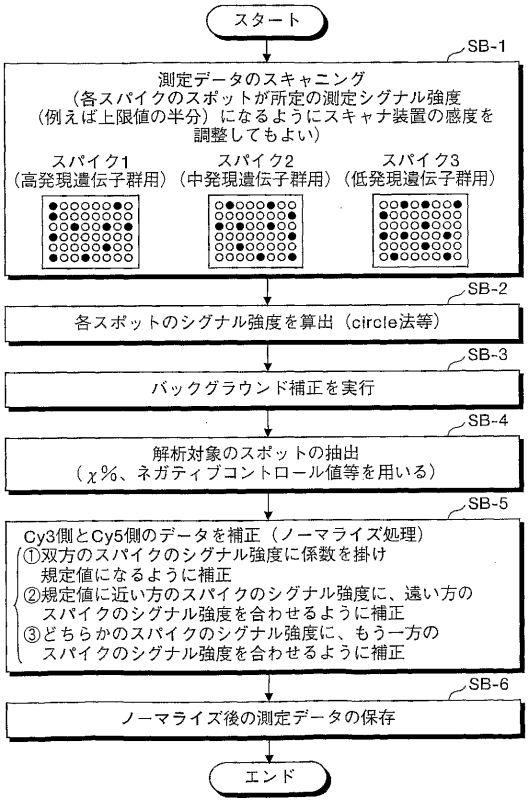

【図２】

第２図



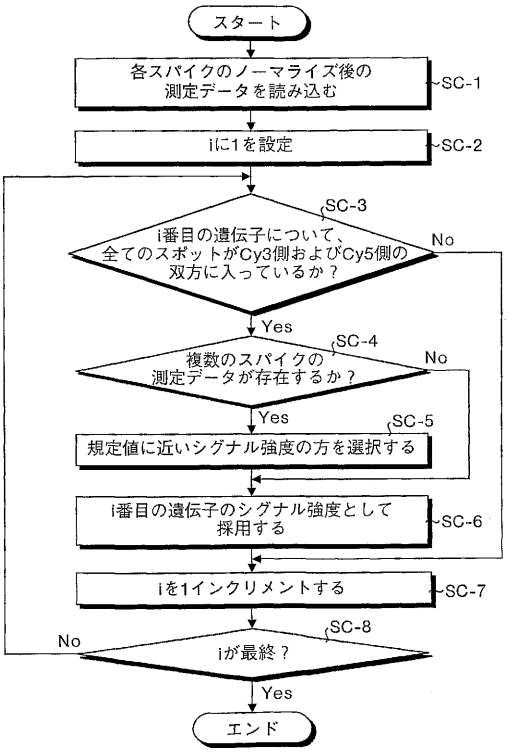

【図３】

第３図



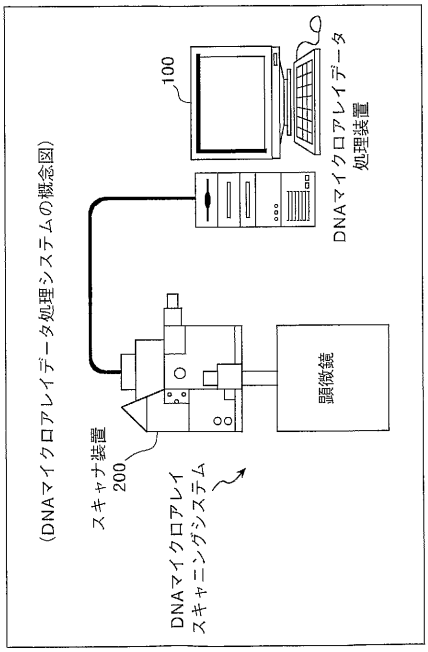

【図４】

第４図



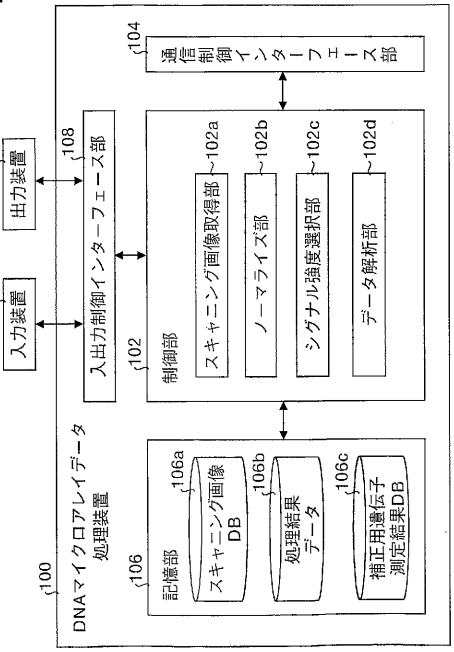
【図５】

第５図



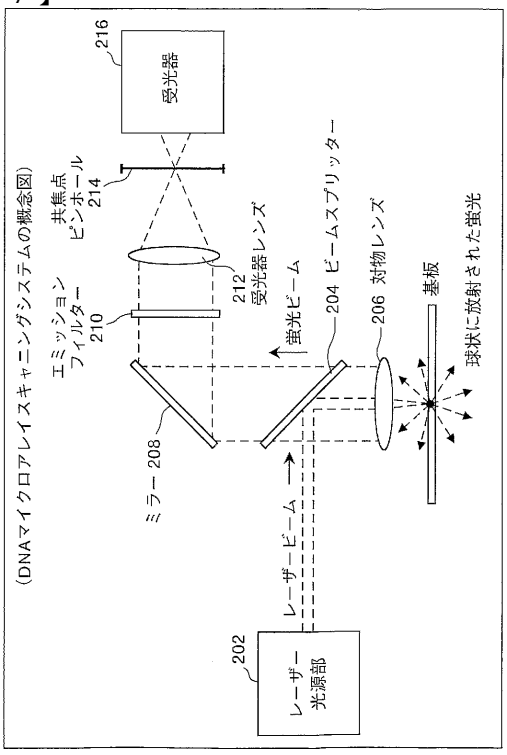

【図６】

第６図



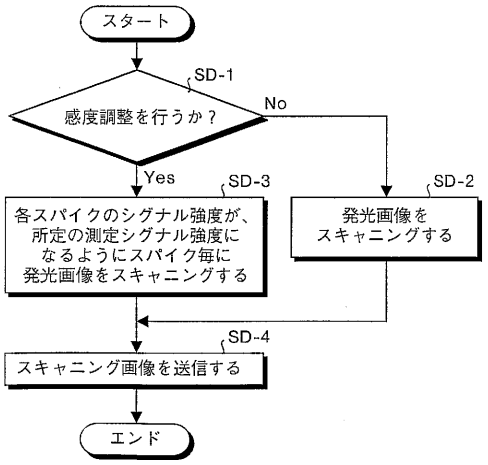

【図７】

第７図



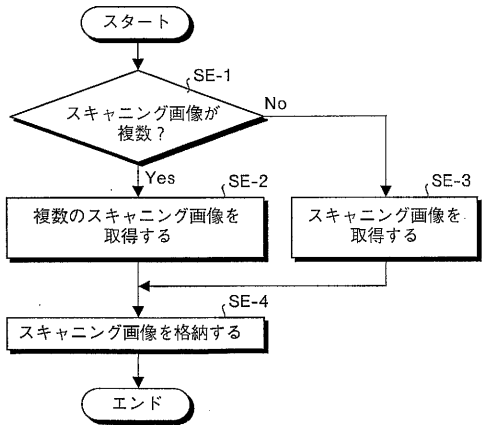

【図８】

第８図



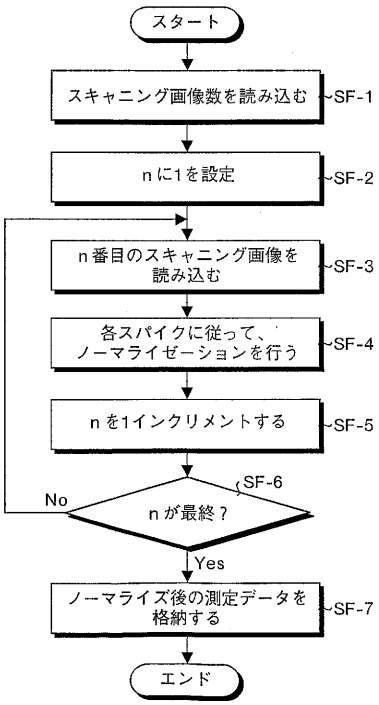
【図９】

第９図



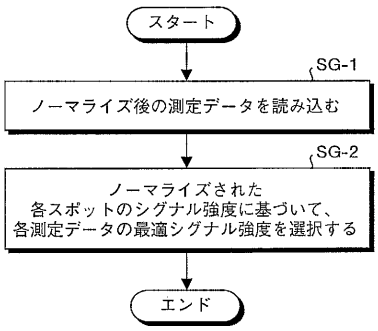

【図１０】

第10図



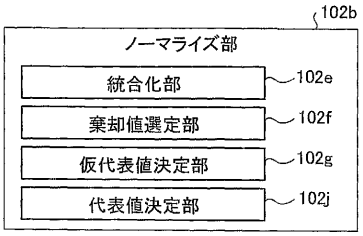

【図１１】

第11図



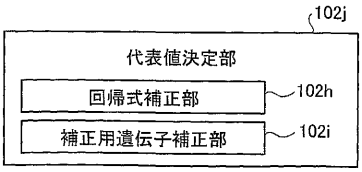

【図１２】

第12図



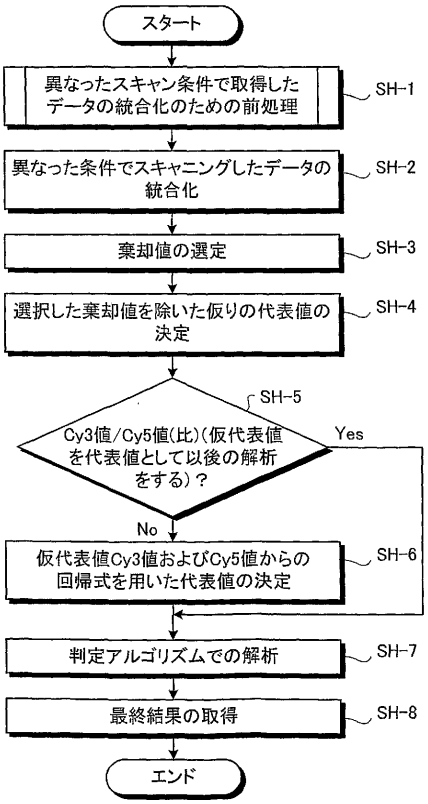

【図１３】

第13図



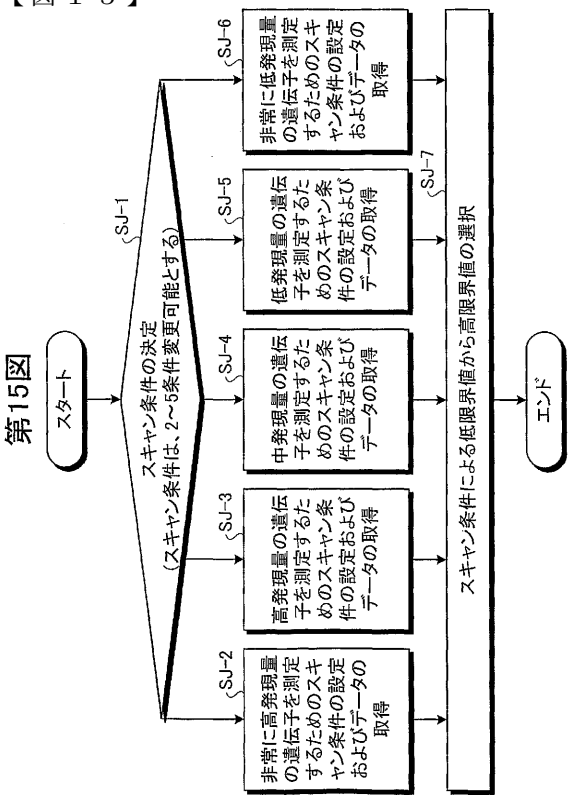
【図１４】

第14図



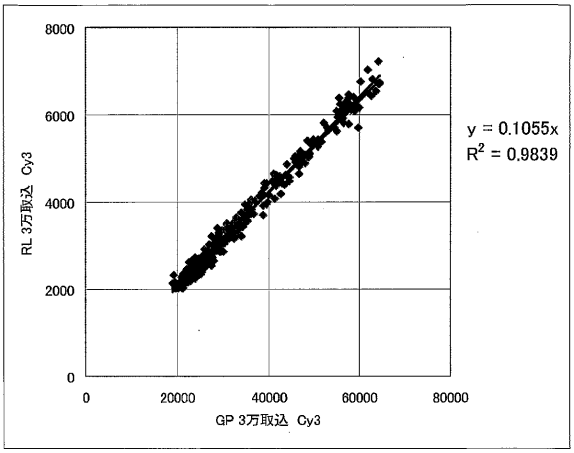

【図１５】

第15図



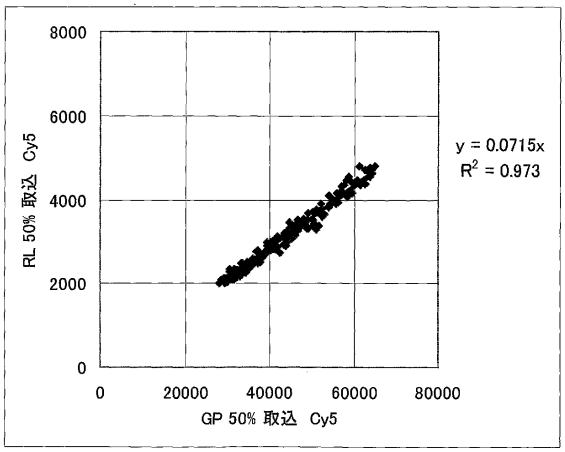

【図１６】

第16図



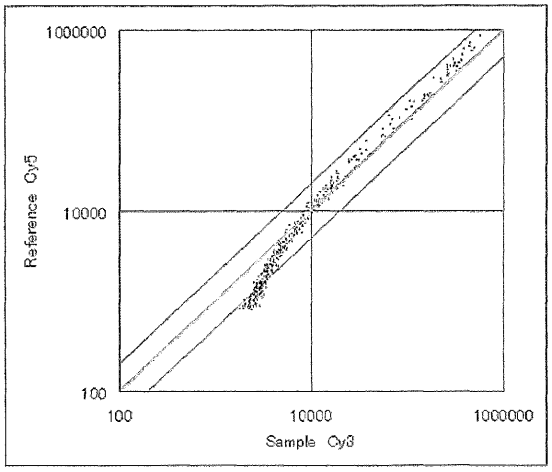

【図１７】

第17図



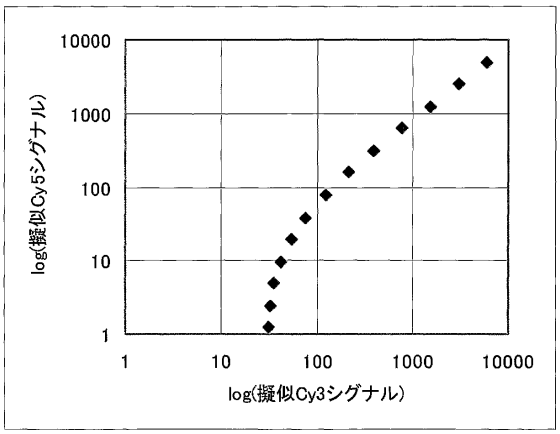
【図１８】

第18図



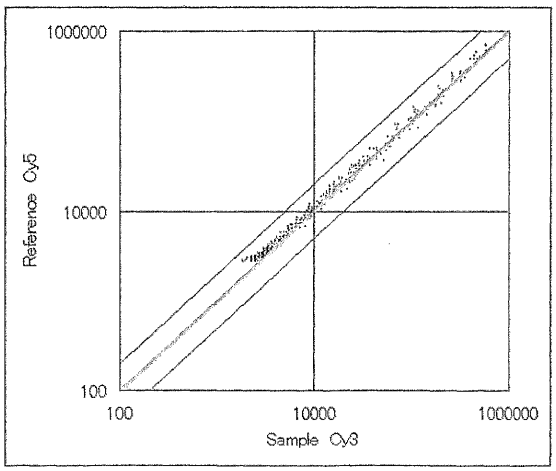

【図１９】

第19図



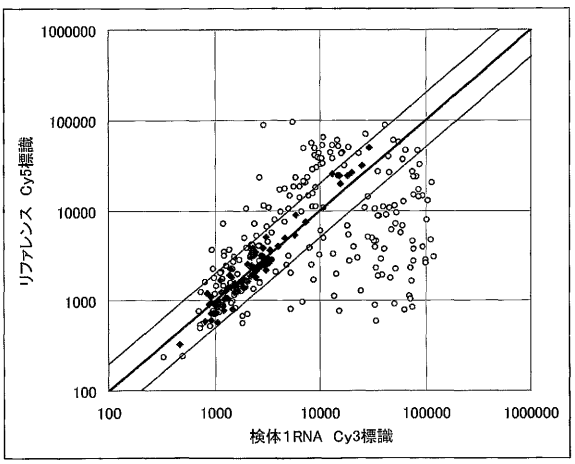

【図２０】

第20図



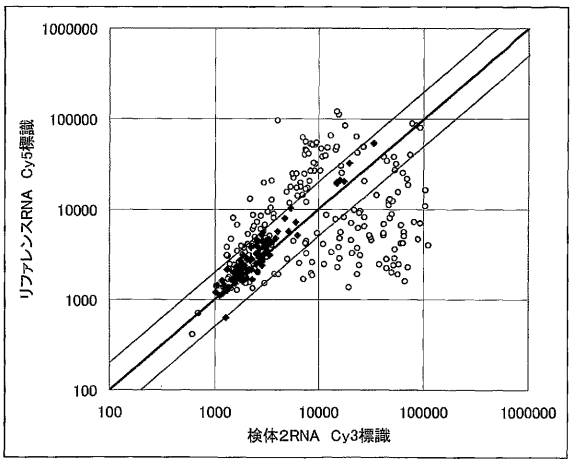

【図２１】

第21図



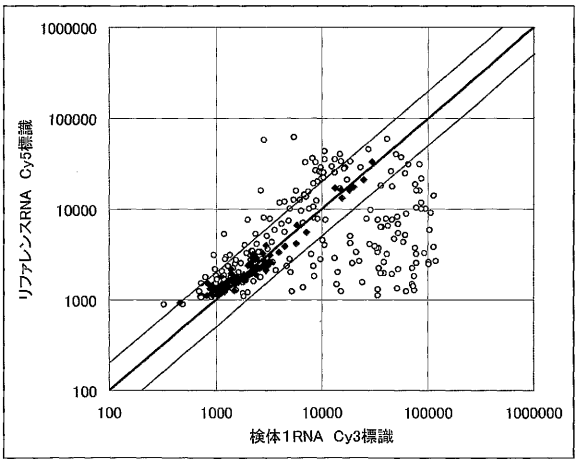
【図２２】

第22図



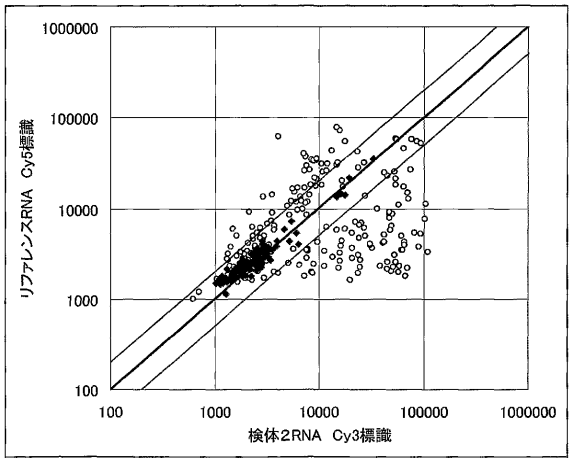

【図２３】

第23図



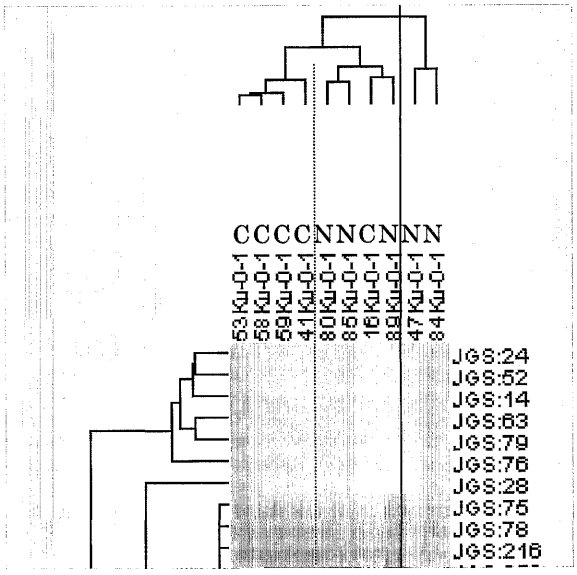

【図２４】

第24図



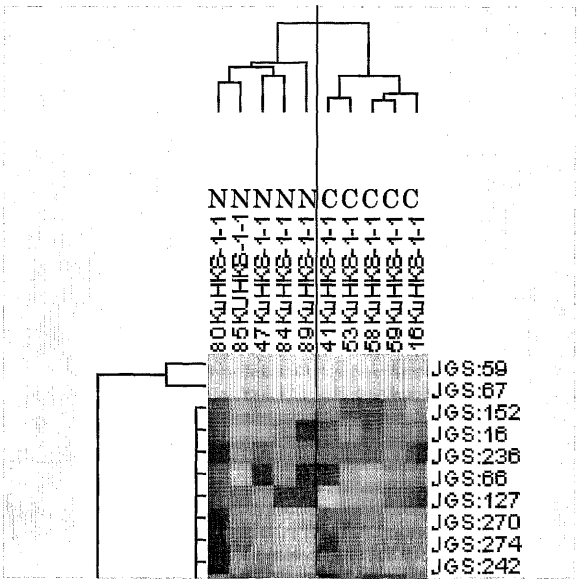

【図２５】

第25図



【図２６】

第26図


NNNNNCCCCC
80KuHKB-1-1
85KUHKB-1-1
47KuHKB-1-1
84KuHKB-1-1
89KuHKB-1-1
41KuHKB-1-1
63KuHKB-1-1
68KuHKB-1-1
59KuHKB-1-1
16KuHKB-1-1
JGS:59
JGS:67
JGS:152
JGS:16
JGS:236
JGS:66
JGS:127
JGS:270
JGS:274
JGS:242

【国際調査報告】

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/06677

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[7] G01N33/53, C12N15/09, C12M1/00, C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[7] G01N33/53, C12N15/09, C12M1/00, C12Q1/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

| | | | |
|---|---|---|---|
| Jitsuyo Shinan Koho | 1922-1996 | Toroku Jitsuyo Shinan Koho | 1994-2003 |
| Kokai Jitsuyo Shinan Koho | 1971-2003 | Jitsuyo Shinan Toroku Koho | 1996-2003 |

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | JP 2001-41892 A (Hitachi Software Engineering Co., Ltd.), 16 February, 2001 (16.02.01), & EP 1116949 A | 1-51 |
| A | WO 01/042512 A (INCYTE GENOMICS, INC.), 14 June, 2001 (14.06.01), & JP 2003-516166 A | 1-51 |
| A | WO 01/040303 A (ISIS INNOVATION LTD.), 07 June, 2001 (07.06.01), & JP 2003-515328 A | 1-51 |

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

| | |
|---|---|
| Date of the actual completion of the international search<br>19 September, 2003 (19.09.03) | Date of mailing of the international search report<br>07 October, 2003 (07.10.03) |
| Name and mailing address of the ISA/<br>Japanese Patent Office<br><br>Facsimile No. | Authorized officer<br><br>Telephone No. |

Form PCT/ISA/210 (second sheet) (July 1998)

国際調査報告

国際出願番号 PCT/JP03/06677

A. 発明の属する分野の分類（国際特許分類（IPC））

Int. Cl.⁷ G01N33/53、C12N15/09, C12M1/00、C12Q1/68

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int. Cl.⁷ G01N33/53、C12N15/09, C12M1/00、C12Q1/68

最小限資料以外の資料で調査を行った分野に含まれるもの

| | |
|---|---|
| 日本国実用新案公報 | 1922－1996年 |
| 日本国公開実用新案公報 | 1971－2003年 |
| 日本国登録実用新案公報 | 1994－2003年 |
| 日本国実用新案登録公報 | 1996－2003年 |

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

| 引用文献のカテゴリー＊ | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示 | 関連する請求の範囲の番号 |
|---|---|---|
| A | JP 2001－41892 A（日立ソフトウエアエンジニアリング株式会社） 2001.02.16 & EP 1116949 A | 1－51 |
| A | WO 01/042512 A (INCYTE GENOMICS, INC.) 2001.06.14 & JP 2003－516166 A | 1－51 |
| A | WO 01/040303 A (ISIS INNOVATION LIMITED) 2001.06.07 & JP 2003－515328 A | 1－51 |

☐ C欄の続きにも文献が列挙されている。 ☐ パテントファミリーに関する別紙を参照。

＊ 引用文献のカテゴリー
「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

| | |
|---|---|
| 国際調査を完了した日 19.09.03 | 国際調査報告の発送日 07.10.03 |
| 国際調査機関の名称及びあて先<br>日本国特許庁（ISA/JP）<br>郵便番号100－8915<br>東京都千代田区霞が関三丁目4番3号 | 特許庁審査官（権限のある職員）<br>亀田 宏之 印 2J 9015<br>電話番号 03－3581－1101 内線 3251 |

様式PCT/ISA/210（第2ページ）（1998年7月）

フロントページの続き

(51)Int.Cl.[7]　　ＦＩ
ＧＯ１Ｎ　37/00　　Ｃ１２Ｎ　15/00　　Ａ

(72)発明者　伊東　篤志
東京都八王子市小宮町５１　株式会社ジェー・ジー・エス内
(72)発明者　伊藤　哲
東京都八王子市小宮町５１　株式会社ジェー・ジー・エス内
(72)発明者　寺崎　浩司
東京都八王子市小宮町５１　株式会社ジェー・ジー・エス内

（注）この公表は、国際事務局（ＷＩＰＯ）により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願（日本語実用新案登録出願）の国際公開の効果は、特許法第１８４条の１０第１項(実用新案法第４８条の１３第２項）により生ずるものであり、本掲載とは関係ありません。